

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: September 10, 2004, 13:26:20 ; Search time 25 Seconds  
(without alignments)  
601.931 Million cell updates/sec

Title: US-09-905-247-1

Perfect score: 1432

Sequence: 1 MRFVLVSLILTLFLTPA.....AAMDDVPMVTQSFQCGSYAL 289

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1422	99.3	289	1	RIP2 TRIKI
2	1377	96.2	289	1	RIPS TRIKI
3	1164.5	81.3	280	1	RIP1 BRYDI
4	859.5	60.0	286	1	RIP1 MOMCH
5	802.5	56.0	277	1	RIPA LUFY
6	756	52.8	286	1	RIP2 MOMBA
7	736	51.4	286	1	RIP1 CUCFI
8	679.5	47.5	294	1	RIP1 TRIAN
9	673	47.0	250	1	RIPB LUFY
10	537.5	37.5	282	1	RIP2 BRYDI
11	432.5	30.2	563	1	NIGB SAMNI
12	419	29.3	576	1	RIC1 RICCO
13	394.5	27.5	564	1	AGGL RICCO
14	343	24.0	562	1	ABRC ABRPR
15	337.5	23.6	527	1	ABRA ABRPR
16	333	23.3	528	1	ABRA ABRPR
17	332	23.2	254	1	MLA3 VISAL
18	322	22.5	254	1	MLA1 VISAL
19	319	22.3	316	1	RIPG GELMU
20	271.5	19.0	294	1	RIPA PHVAM
21	265	18.5	313	1	RIP1 PHVAM
22	237.5	16.6	261	1	RIPS PHVAM
23	215.5	15.0	310	1	RIP2 PHVAM
24	198.5	13.9	278	1	RIPP MIRJA
25	168	11.7	45	1	RIP2 TRIKI
26	168	11.7	299	1	RIP6 SAPOF
27	167	11.7	292	1	RIP7 SAPOF
28	162.5	11.3	253	1	RIP2 SAPOF
29	161.5	11.3	253	1	RIP5 SAPOF
30	152.5	10.6	293	1	RIP0 DIACA
31	138.5	9.7	236	1	RIP3 SAPOF
32	124.5	8.7	319	1	SLTA BP333
33	121	8.4	280	1	RIP2 HORVU

34 118.5 8.3 315 1 SLTA BP319  
35 117.5 8.2 315 1 SLTA BP330  
36 117 8.2 280 1 RIP1 HORVU  
37 102.5 7.2 560 1 JI60 HORVU  
38 97 6.8 454 1 DNAC BACSU  
39 95.5 6.7 504 1 FLIC SALSE  
40 94.5 6.6 504 1 FLIC SALBU  
41 92.5 6.5 939 1 SLAP CAMPE  
42 92 6.4 44 1 RIP3 MOMCH  
43 92 6.4 475 1 MURC WIGBR  
44 91.5 6.4 2748 1 NUM1 YEAST  
45 91 6.4 468 1 HPCC ECOLI

## ALIGNMENTS

RESULT 1  
RIP2 TRIKI  
ID RIP2 TRIKI STANDARD; PRT; 289 AA.  
AC P0989;  
DT 01-MAR-1989 (Rel. 10, Created)  
DT 01-NOV-1990 (Rel. 16, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Ribosome-inactivating protein alpha-trichosanthin precursor  
DE (EC 3.2.2.22) (rRNA N-glycosidase) (Alpha-TCS).  
OS Trichosanthes kirilowii (Mongolian snake-gourd).  
ES Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids 1; Cucurbitales; Cucurbitaceae; Trichosanthes.  
ON NCBI\_TaxID=3677;  
RX MEDLINE=91153657; PubMed=199291;  
RA Shaw P.C., Yung M.H., Zhu R.H., Ho W.K.K., Ng T.B., Yeung H.W.;  
RT "Cloning of trichosanthin cDNA and its expression in *Escherichia coli*.";  
RL Gene 97:267-272(1991).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Maximowicz;  
RX MEDLINE=90256790; PubMed=2341400;  
RA Chow T., Feldman R.A., Lovett M., Piatak M.;  
RT "Isolation and DNA sequence of a gene encoding alpha-trichosanthin, a type I ribosome-inactivating protein.";  
RL J. Biol. Chem. 265:8670-8674 (1990).  
RN [3]  
RP SEQUENCE OF 24-270.  
RC STRAIN=Maximowicz; TISSUE=Tuberos root;  
RX MEDLINE=90256789; PubMed=2341399;  
RA Collins E.J., Robertus J.D., Lopresti M., Stone K.L., Williams K.R., Wu P., Hwang K., Piatak M.;  
RT "Primary amino acid sequence of alpha-trichosanthin and molecular models for abrin A-chain and alpha-trichosanthin.";  
RL J. Biol. Chem. 265:8665-8669 (1990).  
RN [4]  
RP SEQUENCE OF 24-270.  
RC TISSUE=tuberos root;  
RA Wang Y., Qian R.Q., Gu Z.W., Jin S.W., Zhang L.Q., Xia Z.X., Tian G.Y., Ni C.Z.;  
RT "Scientific evaluation of Tian Hua Fen (THF): history, chemistry and application.";  
RL Pure Appl. Chem. 58:789-798 (1986).  
RN [5]  
RP X-RAY CRYSTALLOGRAPHY (1.88 ANGSTROMS).  
RX MEDLINE=94344957; PubMed=8066085;  
RA Zhou F., Fu Z., Chen M., Lin Y., Pan K.;  
RT "Structure of trichosanthin at 1.88-A resolution.";  
RL Proteins 19:4-13(1994).  
RN [6]  
RP X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).  
RX MEDLINE=95344383; PubMed=7619070;

P08026 bacterioph  
P10149 bacterioph  
P22244 hordeum vul  
Q00531 hordeum vul  
P37469 bacillus su  
Q06983 salmonella  
Q06969 salmonella  
P35827 campylobact  
P24817 momordica c  
Q8d227 wigglewort  
Q00402 saccharomyc  
P42269 escherichia

RA Huang Q., Liu S., Tang Y., Jin S., Wang Y.;  
RT "Studies on crystal structures, active-centre geometry and  
RL deprotonating mechanism of two ribosome-inactivating proteins.";  
CC Biochem J. 309:285-298(1995).  
CC -!- FUNCTION: Trichosanthin is an abortion-inducing protein. It is  
CC capable of inhibiting HIV-1 infection and replication. It  
CC inactivates eukaryotic 60S ribosomal subunits.  
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
CC specific adenosine on the 28S rRNA.  
CC -!- SIMILARITY: Belongs to the ribosome-inactivating protein family.  
CC Type 1 RIP subfamily.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: M34858; AAA34207.1; -;  
CC EMBL: J05434; AAA34206.1; -;  
CC PIR: JT0566; RLTT.  
CC PDB: 1MRJ; 07-FEB-95.  
CC PDB: 1MRK; 07-FEB-95.  
CC PDB: 1TCS; 10-JUL-95.  
CC PDB: 1J4G; 28-JAN-03.  
CC PDB: 1NLI; 17-JUN-03.  
CC PDB: 1QB2; 24-APR-00.  
CC InterPro: IPR001574; RIP.  
CC Pfam: PF00161; RIP; 1.  
CC PRINTS: PR00396; SHIGARICIN.  
CC PROSITE: PS00275; SHIGA\_RICIN; 1.  
CC Plant defense; Antiviral; Protein synthesis inhibitor; Hydrolase;  
KW Toxin; Signal; 3D-structure.  
FT SIGNAL 1 23  
FT CHAIN 24 270  
FT  
FT PROPEP 271 289  
FT ACT SITE 183 183  
FT CONFLICT 57 60  
FT CONFLICT 82 84  
FT CONFLICT 87 87  
FT CONFLICT 92 92  
FT CONFLICT 143 144  
FT CONFLICT 196 196  
FT CONFLICT 215 216  
FT CONFLICT 231 231  
FT CONFLICT 234 234  
FT CONFLICT 246 266  
FT CONFLICT 247 247  
FT STRAND 25 28  
FT TURN 30 31  
FT TURN 34 46  
FT HELIX 47 47  
FT TURN 47 47  
FT STRAND 50 54  
FT TURN 55 56  
FT STRAND 57 60  
FT HELIX 66 69  
FT STRAND 70 76  
FT TURN 78 79  
FT STRAND 82 88  
FT TURN 89 92  
FT STRAND 93 99  
FT TURN 100 101  
FT STRAND 102 105  
FT HELIX 109 114  
FT TURN 115 117  
FT TURN 120 121  
FT STRAND 124 127  
FT HELIX 134 141  
FT TURN 142 142  
FT HELIX 145 147

FT STRAND 150 150  
FT HELIX 152 163  
FT TURN 164 185  
FT HELIX 167 180  
FT TURN 181 181  
FT HELIX 182 186  
FT STRAND 187 187  
FT HELIX 188 195  
FT TURN 196 196  
FT STRAND 202 202  
FT HELIX 206 226  
FT TURN 227 230  
FT STRAND 231 239  
FT TURN 241 242  
FT STRAND 245 250  
FT TURN 251 252  
FT HELIX 254 258  
FT TURN 259 259  
FT STRAND 260 260  
FT TURN 263 263  
FT STRAND 266 268  
SQ SEQUENCE 289 AA; 51676 MW; 5CE09BB630575BB9 CRC64;  
Query Match 99.3%; Score 1422; DB 1; Length 289;  
Best Local Similarity 99.3%; Pred. No. 7.9e-105;  
Matches 287; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MIRFLVLSLLILTLFTTPPAVEGDVSRFLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60  
DB 1 MIRFLVLSLLILTLFTTPPAVEGDVSRFLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60  
QY 61 RSLPGSQRYALHILTNVADETISVAIDVTNVIYMGYRAGDTSYFFNGASATEAAKVF 120  
DB 61 RSLPGSQRYALHILTNVADETISVAIDVTNVIYMGYRAGDTSYFFNGASATEAAKVF 120  
QY 121 DAMRKVTLFPGYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMVL 180  
DB 121 DAMRKVTLFPGYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMVL 180  
QY 181 TSEAAARYKFEIQIGKRVKDTLPALAIISLNSWSALSQKQIASTNNGQFSPVVLIN 240  
DB 181 TSEAAARYKFEIQIGKRVKDTLPALAIISLNSWSALSQKQIASTNNGQFSPVVLIN 240  
QY 241 AQNRVTITNDVAGVTSNIALLLNNMAAMDDDDVPMTQSGCGSYAL 289  
DB 241 AQNRVTITNDVAGVTSNIALLLNNMAAMDDDDVPMTQSGCGSYAI 289  
RESULT 2  
RIPS TRIKI STANDARD; PRT; 289 AA.  
AC P24478;  
DT 01-MAR-1992 (Rel. 21, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Ribosome-inactivating protein karasurin precursor (EC 3.2.2.22) (rRNA  
DE N-glycosidase).  
OS Trichosanthin kirilowii (Mongolian snake-gourd).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosid I; Cucurbitales; Cucurbitaceae; Trichosanthes.  
OX NCBI\_TaxID=3677;  
RN [1] SEQUENCE FROM N.A.  
RP TISSUE=Root tuber;  
RC MEDLINE=97356562; PubMed=9212998;  
RX Mizukami H., Iida K., Kondo T., Ogiwara Y.;  
RA "Cloning and bacterial expression of a gene encoding ribosome-  
RT inactivating proteins, karasurin-A and karasurin-C, from Trichosanthes  
RT kirilowii var. japonica";  
RL Biol. Pharm. Bull. 20:711-713(1997).  
RN [2]  
RP SEQUENCE OF 24-270.

```

RX MEDLINE=92005921; PubMed=1914000;
RA Toyokawa S., Takeda T., Kato Y., Wakabayashi K., Ogihara Y.;
RT "The complete amino acid sequence of an abortifacient protein,
RL karasurin.";
RC Chem. Pharm. Bull. 39:1244-1249(1991).
CC -!- FUNCTION: Abortion-inducing protein. It inactivates eukaryotic
CC 60S ribosomal subunits.
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -!- SIMILARITY: Belongs to the ribosome-inactivating protein family.
CC Type 1 RIP subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB000666; BAA21786.1; -
DR PIR; JC5606; JC5606.
DR PIR; J00393; J00393.
DR HSP; P09989; IMRJ.
DR InterPro; IPR001574; RIP.
DR Pfam; PF00161; RIP; 1.
DR PRINTS; PR00396; SHIGARICIN.
DR PROSITE; PS00275; SHIGA_RICIN; 1.
KW Plant defense; Antiviral; Protein synthesis inhibitor; Hydrolase;
KW Toxin; Signal.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 270 KARASURIN-C.
FT CHAIN 24 270 KARASURIN-A.
FT PROPEP 271 289 REMOVED IN MATURE FORM.
FT ACT SITE 183 183 BY SIMILARITY.
SQ SEQUENCE 289 AA; 31704 MW; 883D3E3242887B26 CRC64;

Query Match 96.2%; Score 1377; DB 1; Length 289;
Best Local Similarity 95.8%; Pred No. 2.7e-101;
Matches 277; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 MIRFLVLSLLILTLTPPAVEGVDSFRLSGATSSYGVFISNLRKALPNERKLYDIPLL 60
Db 1 MIRFLVFSLLILTLTPPAVEGVDSFRLSGATSSYGVFISNLRKALPYERKLYDIPLL 60
QY 61 RSLFPGSQRYALHILTNVADETISVAIDVTNYVMYGRAGDTSYFPNGASATEAAKYVEK 120
Db 61 RSTLPGSQRYALHILTNVADETISVAIDVTNYVMYGRAGDTSYFPNEASATEAAKYVEK 120
QY 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPLGLPALDSAITTLFYNNANSAALMVLIOQ 180
Db 121 DAKRKVTLPYSGNYERLQTAAGKIRENIPLGLPALDSAITTLFYNNANSAALMVLIOQ 180
QY 181 TSEAAKYFIEQIGKRVKDTFLPSLAIIISLNSMALSQKIQIASTNNQFESPVLIN 240
Db 181 TSEAAKYFIEQIGKRVKDTFLPSLAIIISLNSMALSQKIQIASTNNQFETPVLIN 240
QY 241 AQQRVTITNDVAGVVTNSIALLLNNNNAAMDDVPMTQSGCGSYAL 289
Db 241 AQQRVTITNDVAGVVTNSIALLLNNNNAAMDDVPMTQSGCGSYAI 289

RESULT 3
RIP1_BRYDI
ID RIP1_BRYDI STANDARD; PRT; 290 AA.
AC P33185; Q98819;
DT 01-OCT-1993 (Rel. 27, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Ribosome-inactivating protein bryodin I precursor (EC 3.2.2.22) (rRNA
DE N-glycosidase) (BD1).
OS Bryonia dioica (Red Bryony).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

```

```

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eutroside I; Cucurbitales; Cucurbitaceae; Bryonia.
OX NCBI_TaxID=3652;
RN [1]
RP SEQUENCE FROM N.A., MUTAGENESIS OF GLU-12, AND X-RAY CRYSTALLOGRAPHY
RP (2.1 ANGSTROMS).
RC TISSUE=Leaf;
RX MEDLINE=97228081; PubMed=9115985;
RA Gawlak S.L., Neubauer M., Klei H.E., Chang C.Y.Y., Einspahr H.M.,
RA Siegall C.B.;
RT "Molecular, biological, and preliminary structural analysis of
RT recombinant bryodin 1, a ribosome-inactivating protein from the plant
RL Bryonia dioica.";
RL Biochemistry 36:3095-3103(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA Siegall C.B.;
RT "Cloning and expression of a gene encoding bryodin 1 from Bryonia
RT dioica.";
RT Patent number US5541110, 30-JUL-1996.
RN [3]
RP SEQUENCE OF 24-66.
RC TISSUE=Seed;
RX MEDLINE=89326691; PubMed=2753596;
RA Montecucchi P.C., Lazzarini A.M., Barbieri L., Stirpe F., Soria M.,
RA Lappi D.;
RT "N-terminal sequence of some ribosome-inactivating proteins.";
RL Int. J. Pept. Protein Res. 33:263-267(1989).
RN [4]
RP SEQUENCE OF 24-43.
RC TISSUE=Root;
RX MEDLINE=95151812; PubMed=7849072;
RA Siegall C.B., Gawlak S.L., Chace D., Wolff E.A., Mixan B.,
RA Marguardt H.;
RT "Characterization of ribosome-inactivating proteins isolated from
RT Bryonia dioica and their utility as carcinoma-reactive
RT immunokonjugates.";
RL Bioconj. Chem. 5:423-429(1994).
CC -!- FUNCTION: Ribosome-inactivating protein of type 1, inhibits
CC protein synthesis in animal cells.
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -!- PTM: Appears to undergo proteolytic cleavage in the C-terminal to
CC produce a shorter protein.
CC -!- BIOTECHNOLOGY: Especially useful as immunotoxin for
CC pharmacological applications as it has low toxicity in rats and
CC mice but is potent once inside target cells.
CC -!- SIMILARITY: Belongs to the ribosome-inactivating protein family.
CC Type 1 RIP subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J24020; -; NOT ANNOTATED_CDS.
DR PIR; S16491; S16491.
DR PDB; 1BRY; 04-MAR-98.
DR InterPro; IPR001574; RIP.
DR Pfam; PF00161; RIP; 1.
DR PRINTS; PR00396; SHIGARICIN.
DR PROSITE; PS00275; SHIGA_RICIN; 1.
KW Plant defense; Protein synthesis inhibitor; Hydrolase; Toxin;
KW 3D-structure; Multigene family; Glycoprotein; Signal.
FT SIGNAL 1 23
FT CHAIN 24 270 RIBOSOME-INACTIVATING PROTEIN BRYODIN 1.
FT PROPEP 271 290 MISSING IN MATURE PROTEIN.
FT ACT SITE 183 183 BY SIMILARITY.
FT ACT SITE 212 212
FT CARBOHYD 214 214 N-LINKED (GLCNAC...) (POTENTIAL).

```

```

FT CARBOHYD 250 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT MUTAGEN 212 E->K: 10-FOLD REDUCTION IN ACTIVITY.
FT CONFLICT 61 RSSIS -> LRHXI (IN REF. 3).
FT STRAND 25 Mismatches 26; Indels 1; Gaps 1;
FT TURN 30 1; Length 290;
FT HELIX 34 46
FT TURN 47 47
FT STRAND 50 54
FT TURN 55 56
FT STRAND 57 60
FT HELIX 66 69
FT STRAND 70 76
FT TURN 78 79
FT STRAND 82 88
FT TURN 89 92
FT STRAND 93 99
FT TURN 100 101
FT STRAND 102 105
FT HELIX 109 114
FT TURN 115 117
FT TURN 120 121
FT STRAND 124 127
FT HELIX 134 141
FT TURN 142 142
FT STRAND 145 147
FT HELIX 150 150
FT STRAND 152 163
FT TURN 164 165
FT HELIX 167 186
FT STRAND 187 187
FT HELIX 188 196
FT STRAND 202 202
FT TURN 206 213
FT HELIX 214 214
FT TURN 215 225
FT STRAND 226 230
FT TURN 231 239
FT STRAND 241 242
FT TURN 245 250
FT STRAND 251 252
FT HELIX 254 257
FT TURN 258 259
FT STRAND 260 260
FT TURN 263 263
FT HELIX 266 268
SQ SEQUENCE 290 AA; 31788 MW; E966CD9C031A42DB CRC64;

Query Match 81.3%; Score 1164.5; DB 1; Length 290;
Best Local Similarity 81.0%; Pred. No. 1.5e-84;
Matches 235; Conservative 26; Mismatches 28; Indels 1; Gaps 1;

QY 1 MIRELVLSLLITLTFTTAVSGDVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60
Db 1 MKLVLLWLLITLTKSTVSGDVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60

QY 61 RSSLPQSORYALIHNTYADETISVAIDVTNVIYIMGYRAGDTSYFFNGASATEAAKYVEK 120
Db 61 RSSISGSGRYTLHLNTYADETISVAIDVTNVIYIMGYLAGDVSYFFNEASATEAAKFVEK 120

QY 121 DAMKVTLPYSGNYERLQTAACKIRENIPGLPALDSAITTLFYNNANSASALMVLIQS 180
Db 121 DAKKVTLPYSGNYERLQTAACKIRENIPGLPALDSAITTLFYNNANSASALMVLIQS 180

QY 181 TSEARYKFIEQIGKRVDKTFLPSLAIISLNSWSALSCKIQIASTNNNGOPESPVLIN 240
Db 181 TAESARYKFIEQIGKRVDKTFLPSLAIISLNSWSALSCKIQIASTNNNGOPESPVLID 240

QY 241 AQNRVTITNDAGVVTGNIALLLNRNNMAAMDDVPMQ-SFGCGSYAL 289
Db 241 GNNQRVSITNASARVVTGNIALLLNRNNIAAGEDISMTLIGFERGLYGI 290

```

.RESULT 4

```

RIPI1 MOMCH
ID RIPI1 MOMCH STANDARD; PRT; 286 AA.
AC P16094; P24697;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Ribosome-inactivating protein momordin I precursor (EC 3.2.2.22) (rRNA
DE N-glycosidase) (Alpha-momorcharin) (Alpha-MMC)
OS Momordica charantia (Bitter melon) (Balsam pear).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids I; Cucurbitales; Cucurbitaceae; Momordica.
OC NCBI_TaxID=3673;
RN [1]_TaxID=3673;
RP SEQUENCE FROM N.A.
RC TISSUE=Seed;
RX MEDLINE=91159486; PubMed=2001404;
RA Ho W.K.K., Liu S.C., Shaw P.C., Yeung H.W., Ng T.B., Chan W.Y.;
RT "Cloning of the cDNA of alpha-momorcharin: a ribosome inactivating
RT protein."
RL Biochim. Biophys. Acta 1088:311-314(1991).
RN [2]
RP SEQUENCE OF 24-38.
RC TISSUE=Seed;
RX MEDLINE=89326691; PubMed=2753596;
RA Montecucchi P.C., Lazzarini A.M., Barbieri L., Stirpe F., Soria M.,
RA Lappi D.;
RT "N-terminal sequence of some ribosome-inactivating proteins."
RL Int. J. Pept. Protein Res. 33:263-267(1989).
RN [3]
RP SEQUENCE OF 24-70.
RC TISSUE=Seed;
RX MEDLINE=89005108; PubMed=3262509;
RA Casellas P., Dussosoy D., Falasca A.I., Barbieri L.,
RA Guillemot J.C., Ferrara P., Bolognesi A., Cennini P., Stirpe F.;
RT Trichokirin, a ribosome-inactivating protein from the seeds of
RT Trichosanthes kirilowii Maximowicz. Purification, partial
RT characterization and use for preparation of immunotoxins."
RL Eur. J. Biochem. 176:581-588(1988).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
RX MEDLINE=94356447; PubMed=8075985;
RA Ren J., Wang Y., Dong Y., Stuart D.I.;
RT "The N-glycosidase mechanism of ribosome-inactivating proteins
RT implied by crystal structures of alpha-momorcharin."
RL Structure 2:7-16(1994).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.16 ANGSTROMS).
RX MEDLINE=94192822; PubMed=8143869;
RA Husain J., Tickle I.J., Wood S.P.;
RT "Crystal structure of momordin, a type I ribosome inactivating
RT protein from the seeds of Momordica charantia."
RL FEBS Lett. 342:154-158(1994).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX MEDLINE=95344383; PubMed=7619070;
RA Huang Q., Liu S., Tang Y., Jin S., Wang Y.;
RT "Studies on crystal structures, active-centre geometry and
RT depurinating mechanism of two ribosome-inactivating proteins."
RL Biochem. J. 309:285-298(1995).
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -!- SIMILARITY: Belongs to the ribosome-inactivating protein family.
CC Type 1 Rlp subfamily.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----

```







Db 1 MNRFSVLMCLVILSLFPHGVPAGDVTKFSLGSHKHSYSKFITSMNALPNACDIYNI 60  
 QY 58 PLLRSSLPQSORYALIHITNVADETISVAIDVTNVMYMGYRAGDTSYFNGASATEAAKY 117  
 Db 61 PLLVPSISGSRRIYLMQLSNVEGNTIMWADVNTVMYMGYLVNGTSYFNETDAQLASKF 120  
 QY 118 VFQ-OTKSTILPYSNGYKQSVARKEKDSIPLGFMALDSALSTUYDYDSRSPAFILVL 179  
 Db 121 VFQ-OTKSTILPYSNGYKQSVARKEKDSIPLGFMALDSALSTUYDYDSRSPAFILVL 179  
 QY 178 IQSTSEAAARYKFEIOQIGKRVDKTFLPSLAISLNSWSALSQKIQIASTNNQGFESPVV 237  
 Db 180 IOTTAAARYKYIEQIIDRISVSKVPDLAALSLENWSLLSKQIQIAKSNNGQFQTPVK 239  
 QY 238 LINAQQRVTITNDVAGVTSNIALLNRRNNMAAMDDVPMT 279  
 Db 240 IINDKXGILTEVNTVSSLVVTKNIMLLNKLNTASPEDHVIST 281

## RESULT 8

RIP1 TRIAN  
 ID RIPI TRIAN STANDARD; PRT; 294 AA.  
 AC P56626; Q9ZQY7;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Type 1 ribosome-inactivating protein trichoanguina precursor  
 DE (EC 3.2.2.22) (rRNA N-glycosidase) (RIP) (Trichoanguin).  
 GN TCA.  
 OS Trichosanthes anguina (Snake gourd).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.  
 OX NCBI\_TaxID=50544;  
 RN [1]  
 RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND CHARACTERIZATION.  
 RC STRAIN=cv. Anguina; TISSUE=Seed;  
 RX MEDLINE=99132006; PubMed=9931318;  
 RA Chow L.-P., Chou M.-H., Ho C.-Y., Chuang C.-C., Pan F.-M., Wu S.-H.,  
 Lin J.-Y.;  
 RT "Purification, characterization and molecular cloning of trichoanguin,  
 a novel type I ribosome-inactivating protein from the seeds of  
 Trichosanthes anguina".  
 RL Biochem. J. 338:211-219 (1999).  
 RN [2]  
 RP SEQUENCE OF 20-264.  
 RC TISSUE=Seed;  
 RA Chow L.-P., Kamo M., Lin J.-Y., Wang S.-H., Ueno Y., Tsugita A.;  
 RT "Amino acid sequence of trichoanguina, a ribosomal-inactivating  
 protein from Trichosanthes anguina seeds".  
 RL J. Biomed. Sci. 3:178-186 (1996).  
 CC -!- FUNCTION: Inhibits protein synthesis by depurinating 28S rRNA in  
 ribosomes.  
 CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
 specific adenosine on the 28S rRNA.  
 CC -!- SIMILARITY: Belongs to the ribosome-inactivating protein family.  
 CC Type 1 RIP subfamily.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AF055086; RAD02686.1; -.  
 CC HSSP; P33185; LBRY.  
 CC InterPro; IPR001574; RIP.  
 CC Pfam; PF00161; RIP; 1.  
 CC PRINTS; PR00396; SHIGARICIN.  
 CC PROSITE; PS00275; SHIGA\_RICIN; FALSE\_NEG.

KW Plant defense; Protein synthesis inhibitor; Hydrolase; Toxin;  
 KW Glycoprotein; Signal.  
 FT SIGNAL 1 19  
 FT CHAIN 20 264  
 TYPE I RIBOSOME-INACTIVATING PROTEIN  
 TRICHOANGUINA.  
 FT PROPEP 265 294  
 FT ACT\_SITE 177 177  
 FT ACT\_SITE 180 180  
 FT CARBOHYD 70 70  
 FT CARBOHYD 220 220  
 FT CONFLICT 51 51  
 FT CONFLICT 65 65  
 FT CONFLICT 84 84  
 FT CONFLICT 152 152  
 FT CONFLICT 174 174  
 FT CONFLICT 245 245  
 FT CONFLICT 294 294  
 FT SEQUENCE 294 AA; 32234 MW; DA4F8B7CE3290994 CRC64;  
 Query Match 47.5%; Score 679.5; DB 1; Length 294;  
 Best Local Similarity 52.2%; Pred. No. 2.3e-46;  
 Matches 144; Conservative 38; Mismatches 91; Indels 3; Gaps 3;  
 QY 5 IVLSSLILTLFTTAVEGDVSFRLSGATSSSYGVFISNLKALPNERKLYDIPILRSSL 64  
 Db 1 MALSPFFLAISLGSPALGDVSFDLSLTATKXSYSSFTQLRDALPTQGTVCIGILLPSTA 60  
 QY 65 PGSQRYALIHITNVADETISVAIDVTNVMYMGYRAGDTSYFNGASATEAAKYVFKDAMR 124  
 Db 61 SGSQWFRFNLNTNDETTVAVNVNTVYI VAYRADVSFFEDTPA-EAFKLIFA-GTK 118  
 QY 125 KVTLPYSNGYERLQTAAGKIRENIPGLPALDSAITTLFFYNANSASALMVLIOSTSEA 184  
 Db 119 TVKLPYSNGYDKLQSVGKQDMIELGIPALSSAITNNVYDYQSTAAALLVLIQTAEA 178  
 QY 185 ARYKFEIOQIGKRVDKTFLPSLAISLNSWSALSQKIQIAS-TNNGQFESPVLINAQ 243  
 Db 179 ARYKFEIOQVSSHSSNFYNGQAVISLENKMGALSQKIQIANRTGHGQFENPVELYNDG 238  
 QY 244 QRVITITNDVAGVTSNIALLNRRNNMAAMDDVPMT 279  
 Db 239 TRFSVTNTSAGVVGKGNIKLLLYKASGVSEYDIPTT 274  
 RESULT 9  
 RIPB LUFYCY  
 ID RIPB LUFYCY STANDARD; PRT; 250 AA.  
 AC P22851;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Ribosome-inactivating protein luffin-B (EC 3.2.2.22) (rRNA N-  
 glycosidase).  
 DE Luffa cylindrica (Smooth loofah) (Sponge gourd).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids I; Cucurbitales; Cucurbitaceae; Luffa.  
 OX NCBI\_TaxID=3670;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Seed;  
 RX MEDLINE=91248488; PubMed=1368666;  
 RA Islam M.R., Hirayama H., Funatsu G.;  
 RT "Complete amino acid sequence of luffin-b, a ribosome-inactivating  
 protein from sponge gourd (luffa cylindrica) seeds.";  
 RL Agric. Biol. Chem. 55:229-238 (1991).  
 CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
 specific adenosine on the 28S rRNA.  
 CC -!- SIMILARITY: Belongs to the ribosome-inactivating protein family.  
 CC Type 1 RIP subfamily.  
 CC PIR; JN0108; JN0108.  
 CC HSSP; PL6094; LAHC.  
 CC InterPro; IPR001574; RIP.  
 CC Pfam; PF00161; RIP; 1.





RA Olsnes S., Kozlov J.V.;  
RT "Ricin";  
RL Toxicon 39:1723-1728(2001).  
RN [8]  
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).  
RX MEDLINE=87165983; PubMed=3558397;  
RA Montfort W., Villafraña J.E., Monzingo A.F., Ernst S.R., Katzin B.,  
RA Rutenber E., Kuong N.H., Hamlin R., Robertus J.D.;  
RT "The three-dimensional structure of ricin at 2.8 Å";  
RL J. Biol. Chem. 262:5398-5403(1987).  
RN [9]  
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF A-CHAIN.  
RX MEDLINE=91352004; PubMed=1881881;  
RA Katzin B.J., Collins E.J., Robertus J.D.;  
RT "Structure of ricin A-chain at 2.5 Å";  
RL Proteins 10:251-259(1991).  
RN [10]  
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF B-CHAIN.  
RX MEDLINE=91352005; PubMed=1881882;  
RA Rutenber E., Robertus J.D.;  
RT "Structure of ricin B-chain at 2.5-Å resolution.";  
RL Proteins 10:260-269(1991).  
RN [11]  
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF A-CHAIN.  
RX MEDLINE=95082010; PubMed=7990130;  
RA Weston S.A., Tucker A.D., Thatcher D.R., Derbyshire D.J.,  
RA Pauptit R.A.;  
RT "X-ray structure of recombinant ricin A-chain at 1.8-Å resolution.";  
RL J. Mol. Biol. 244:410-422(1994).  
RN [12]  
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF MUTANT HIS-215.  
RX MEDLINE=96374222; PubMed=8780513.  
RA Day P.J., Ernst S.R., Frankel A.E., Monzingo A.F., Pascal J.M.,  
RA Molina-Svinth M.C., Robertus J.D.;  
RT "Structure and activity of an active site substitution of ricin A chain";  
RL Biochemistry 35:11098-11103(1996).  
RN [13]  
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF A-CHAIN.  
RX MEDLINE=97240820; PubMed=9086280;  
RA Yan X., Hollis T., Svinth M., Day P., Monzingo A.F., Milne G.W.,  
RA Robertus J.D.;  
RT "Structure-based identification of a ricin inhibitor.";  
RL J. Mol. Biol. 266:1043-1049(1997).  
RN [14]  
RP MUTAGENESIS.  
RX MEDLINE=93165632; PubMed=1287657;  
RA Kin Y., Robertus J.D.;  
RT "Analysis of several key active site residues of ricin A chain by mutagenesis and X-ray crystallography.";  
RL Protein Eng. 5:775-779(1992).  
RN [15]  
RP MUTAGENESIS OF ARG-83; LEU-109; ASP-110; VAL-111 AND ASN-132.  
RX MEDLINE=2252669; PubMed=12627168;  
RA Smallshaw J.E., Ghetie V., Rizo J., Fulmer J.R., Trahan L.L.,  
RA Ghetie M.-A., Vitetta E.S.;  
RT "Genetic engineering of an immunotoxin to eliminate pulmonary vascular leak in mice.";  
RL Nat. Biotechnol. 21:387-391(2003).  
CC -!- FUNCTION: Ricin is highly toxic to animal cells and to a lesser extent to plant cells. The A chain acts as a glycosidase that removes a specific adenine residue from an exposed loop of the 28S rRNA (A4324 in mammals), leading to rRNA breakage. As this loop is involved in elongation factor binding, modified ribosomes are catalytically inactive and unable to support protein synthesis. The A chain can inactivate a few thousand ribosomes per minute, faster than the cell can make new ones. Therefore a single A chain molecule can kill an animal cell. The B chain binds to beta-D-galactopyranoside moieties on cell surface glycoproteins and glycolipids and facilitates the entry into the cell of the A chain; B chains are also responsible for cell agglutination (Lectin activity).  
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one specific adenosine on the 28S rRNA.  
CC -!- SUBUNIT: Disulfide-linked dimer of A and B chains.  
CC -!- DOMAIN: The B chain is composed of two domains, each domain consists of 3 homologous subdomains (alpha, beta, gamma).  
CC -!- BIOTECHNOLOGY: A deglycosylated A chain may be linked to monoclonal antibodies to produce immunotoxins exploited in cancer treatment. However, a point mutation should be introduced to eliminate vascular leak syndrome, a side effect resulting from endothelial damage.  
CC -!- SIMILARITY: In the N-terminal section; belongs to the ribosome-inactivating protein family. Type 2 RIP subfamily.  
CC -!- SIMILARITY: Contains 2 ricin B-type lectin domains.  
CC -!- CAUTION: Ref.4 and Ref.5 sequences have a high number of conflicts with the sequence translated from DNA (Ref.1, Ref.2 and Ref.3).  
CC -!- DATABASE: NAME-Protein Spotlight;  
CC NOTE=Issue 31 of February 2003;  
CC WWW="http://www.expasy.org/spotlight/articles/sptlt031.html".  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; X03179; CAA26939.1; -;  
DR EMBL; X52908; CAA37095.1; -;  
DR EMBL; X02388; CAA26230.1; -;  
DR EMBL; A12892; CAA01058.1; -;  
DR PIR; A24041; RLCSO.  
DR PDB; 2AAI; 31-JAN-94.  
DR PDB; 1APG; 31-JAN-94.  
DR PDB; 1FMP; 31-OCT-93.  
DR PDB; 1IFS; 14-JAN-98.  
DR PDB; 1IFT; 14-JAN-98.  
DR PDB; 1IFU; 14-JAN-98.  
DR PDB; 1R7C; 31-OCT-93.  
DR PDB; 1OBS; 16-JUN-97.  
DR PDB; 1OBT; 16-JUN-97.  
DR PDB; 1BR5; 02-SEP-98.  
DR PDB; 1BR6; 02-SEP-98.  
DR PDB; 1IL3; 16-JAN-02.  
DR PDB; 1IL4; 16-JAN-02.  
DR PDB; 1IL9; 16-JAN-02.  
DR GlycoSuiteDB; P02879;  
DR InterPro; IPR000772; Ricin\_B\_lectin.  
DR InterPro; IPR008997; RicinB\_like.  
DR InterPro; IPR001574; RIP.  
DR Pfam; PF00652; Ricin\_B\_lectin; 6.  
DR Pfam; PF00161; RIP; 1.  
DR PRINTS; PR00396; SHIGARICIN.  
DR SMART; SM00458; RICIN; 2.  
DR PROSITE; PS00231; RICIN\_B\_LECTIN; 2.  
DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
KW Plant defense; Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;  
KW Glycoprotein; Lectin; Signal; 3D-structure.  
FT SIGNAL 1 35  
FT CHAIN 36 302  
FT PEPTIDE 303 314  
FT CHAIN 315 576  
FT DOMAIN 321 448  
FT DOMAIN 451 575  
FT REPEAT 331 373  
FT REPEAT 374 414  
FT REPEAT 417 449  
FT REPEAT 462 497  
FT REPEAT 501 540  
FT REPEAT 543 570  
FT ACT\_SITE 212 212  
FT DISULFID 294 318  
FT DISULFID 334 353  
FT DISULFID 377 394





```

RESULT 14
ABRC ABRPR
ID ABRP ABRPR STANDARD; PRT; 562 AA.
AC P28590;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE ABRin-c precursor [Contains: ABRin-c A chain (EC 3.2.2.22) (rRNA N-glycosidase); ABRin-c B chain].
OS ABRus precatorius (Indian licorice) (Crab's eye).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.
OX NCBI_TaxID=3816;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaf;
RX MEDLINE=931266957; PubMed=2050149;
RA Wood K.A., Lord J.M., Wawrzyniak E.J., Piatak M.;
RT "preproabrin: genomic cloning, characterisation and the expression of
RT the A-chain in Escherichia coli.";
RL Eur. J. Biochem. 198;723-732(1991).
CC -!- FUNCTION: The A chain is responsible for inhibiting protein
CC synthesis through the catalytic inactivation of 60S ribosomal
CC subunits by removing adenine from position 4,324 of 28S rRNA.
CC ABRin-a is more toxic than ricin.
CC -!- FUNCTION: The B chain is a galactose-specific lectin that
CC facilitates the binding of abrin to the cell membrane that
CC precedes endocytosis.
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -!- SUBUNIT: Disulfide-linked dimer of A and B chains.
CC -!- DOMAIN: The B chain is composed of two domains, each domain
CC consists of 3 homologous subdomains (alpha, beta, gamma).
CC -!- SIMILARITY: In the N-terminal section; belongs to the ribosome-
CC inactivating protein family. Type 2 Rip subfamily.
CC -!- SIMILARITY: Contains 2 ricin B-type lectin domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X55667; CA39202.1;
CC PIR; S16022; S16022.
CC HSP; F11140; IABR.
CC InterPro; IPR000772; Ricin_B_lectin.
CC InterPro; IPR008997; RicinB_like.
CC InterPro; IPR001574; RIP.
CC Pfam; PF00652; Ricin_B_lectin; 6.
CC Pfam; PF00161; RIP; I_lectin; 6.
CC PRINTS; PR00396; SHIGARICIN.
CC SMART; SM00458; RICIN; 2.
CC PROSITE; PS00231; RICIN_B_LECTIN; 2.
CC PROSITE; PS00275; SHIGA_RICIN; 1.
CC Plant defense; Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;
CC Glycoprotein; Lectin; Signal; Pyrrolidone carboxylic acid.
CC SIGNAL
CC CHAIN 1 34 BY SIMILARITY.
CC CHAIN 35 285 ABRIN C A CHAIN (BY SIMILARITY).
CC PEPTIDE 286 295 LINKER PEPTIDE (BY SIMILARITY).
CC CHAIN 296 562 ABRIN C B CHAIN (BY SIMILARITY).
CC DOMAIN 307 434 RICIN B-TYPE LECTIN 1.
CC DOMAIN 437 561 RICIN B-TYPE LECTIN 2.
CC REPEAT 317 359 1-ALPHA.
CC REPEAT 360 400 1-BETA.
CC REPEAT 403 435 1-GAMMA.
CC REPEAT 448 483 2-ALPHA.
CC REPEAT 487 526 2-BETA.
CC REPEAT 529 562 2-GAMMA.
CC ACT_SITE 198 BY SIMILARITY.
CC

```

---

```

FT DISULFID 281 303 INTERCHAIN (BY SIMILARITY).
FT DISULFID 320 339 BY SIMILARITY.
FT DISULFID 363 380 BY SIMILARITY.
FT DISULFID 451 464 BY SIMILARITY.
FT DISULFID 490 507 BY SIMILARITY.
FT MOD_RES 35 PYRROLIDONE CARBOXYLIC ACID
(BY SIMILARITY).
FT CARBOHYD 234 234 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 395 395 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 435 435 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 562 AA; 62817 MW; 1FD0ABC7D7BA6278 CRC64;
Query Match 24.0%; Score 343; DB 1; Length 562;
Best Local Similarity 36.4%; Pred. No. 1.6e-19;
Matches 102; Conservative 43; Mismatches 107; Indels 28; Gaps 10;
QY 7 LSLILITLPLT-----PAVEGD-----VSFRLSGATSSSYGVFTSNLRKALPN 50
DB 5 LKLLILCLAWTCFSALRCAARTYPPVATNQDVIKFTTEGATSSYKOFIEALQRUTG 64
QY 51 ERKYLDIPLL--RSSLPQSQRVALIHLTNVADETISVAIDVTNVYIMGYRAGDTSYFFNG 108
DB 65 -GLIHDIPLVDPPTVEERNRYITVELSNSRESIEVGIDVTNAYVAYRAGSQSYFLRD 123
QY 109 ASATEAAKYVFKDAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTL--FYNN 166
DB 124 APAS-ASTYLPFGTQR-YSLRFDGSGYGLERWAHQTRREISLGLQALTHAISFLRSGASN 181
QY 167 ANSASALMVLIOSTSEAAKYKFTQQIGK--RVDKTLPLSLAIISLNSALSQKQIQI 224
DB 182 DEEKARTLIVTIQMASEAARYISNRVGSIRGTAFQPPAPMLSLNNNDNLSSGGVQ- 240
QY 225 ASTNNGQESPVLINAGNQRTVITNDAGVVTNSIALLL 264
DB 241 -QSVQDTFNNVILLSSINRQPVVDLSLHPV-AVLALML 278

```

---

```

RESULT 15
ABRB ABRPR STANDARD; PRT; 527 AA.
AC Q06077; P81374;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE ABRin-b precursor [Contains: ABRin-b A chain (EC 3.2.2.22) (rRNA N-glycosidase); ABRin-b B chain].
OS ABRus precatorius (Indian licorice) (Crab's eye).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.
OX NCBI_TaxID=3816;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93132798; PubMed=7763423;
RA Hung C.-H., Lee M.-C., Lin J.-Y.;
RT "Primary structure of three distinct isoabrin determined by cDNA
RT sequencing. Conservation and significance.";
RL J. Mol. Biol. 229:263-267(1993).
RN [2]
RP SEQUENCE OF 260-527.
RC TISSUE=Seed;
RX MEDLINE=93169023; PubMed=7763422;
RA Kimura M., Sumizawa T., Funatsu G.;
RT "The complete amino acid sequences of the B-chains of abrin-a and
RT abrin-b, toxic proteins from the seeds of Abrus precatorius.";
RL Biosci. Biotechnol. Biochem. 57:166-169(1993).
CC -!- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN
CC SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL
CC SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RRNA.
CC ABRIN-A IS MORE TOXIC THAN RICIN.
CC -!- FUNCTION: The B chain is a galactose-specific lectin that
CC facilitates the binding of abrin to the cell membrane that
CC precedes endocytosis.
CC

```







GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 13:56:41 ; Search time 33 Seconds  
(without alignments)  
452.118 Million cell updates/sec

Title: US-09-905-247-1

Perfect score: 1432

Sequence: 1 MIRFLVLSLILTLFLTPA.....AAMDDDDVPMQTSGCGSYAL 289

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

- 1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/iaa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	1413	98.7	289	1 US-07-923-692C-4	Sequence 4, Appli
2	1413	98.7	289	1 US-08-184-237-4	Sequence 4, Appli
3	1413	98.7	289	1 US-08-482-920-4	Sequence 4, Appli
4	1413	98.7	289	3 US-08-484-341-4	Sequence 4, Appli
5	1413	98.7	289	3 US-08-483-502-4	Sequence 4, Appli
6	1413	98.7	289	4 US-09-726-651A-4	Sequence 4, Appli
7	1321	92.2	267	1 US-08-378-761A-74	Sequence 74, Appl
8	1321	92.2	267	1 US-08-485-286-74	Sequence 74, Appl
9	1202	83.9	247	1 US-08-488-113B-6	Sequence 6, Appli
10	1202	83.9	247	1 US-08-477-484B-6	Sequence 6, Appli
11	1202	83.9	247	2 US-08-646-360-6	Sequence 6, Appli
12	1202	83.9	247	3 US-08-839-765-6	Sequence 6, Appli
13	1202	83.9	247	3 US-09-136-389-6	Sequence 6, Appli
14	1202	83.9	247	3 US-09-610-838-6	Sequence 6, Appli
15	1202	83.9	247	4 US-09-711-485-6	Sequence 6, Appli
16	1164.5	81.3	290	1 US-08-245-754A-2	Sequence 2, Appli
17	1164.5	81.3	290	2 US-08-597-731-2	Sequence 2, Appli
18	1130	78.9	255	1 US-07-901-707-6	Sequence 6, Appli
19	1130	78.9	255	1 US-07-988-430-6	Sequence 6, Appli
20	1130	78.9	255	1 US-08-423-336-6	Sequence 6, Appli
21	1130	78.9	255	5 PCT-US92-09487-6	Sequence 6, Appli
22	1074	75.0	496	3 US-08-902-486-15	Sequence 15, Appli
23	1064	74.3	248	3 US-08-902-486-7	Sequence 7, Appli
24	806.5	56.3	263	1 US-07-901-707-7	Sequence 7, Appli
25	806.5	56.3	263	1 US-07-988-430-7	Sequence 7, Appli
26	806.5	56.3	263	1 US-08-425-336-7	Sequence 7, Appli
27	806.5	56.3	263	1 US-08-488-113B-7	Sequence 7, Appli

28	806.5	56.3	263	1 US-08-477-484B-7	Sequence 7, Appli
29	806.5	56.3	263	2 US-08-646-360-7	Sequence 7, Appli
30	806.5	56.3	263	3 US-08-839-765-7	Sequence 7, Appli
31	806.5	56.3	263	3 US-09-136-389-7	Sequence 7, Appli
32	806.5	56.3	263	4 US-09-610-838-7	Sequence 7, Appli
33	806.5	56.3	263	4 US-09-711-485-7	Sequence 7, Appli
34	806.5	56.3	263	5 PCT-US92-09487-7	Sequence 7, Appli
35	760	53.1	250	1 US-08-378-761A-78	Sequence 78, Appli
36	760	53.1	250	1 US-08-485-286-78	Sequence 78, Appli
37	751	52.4	286	1 US-08-324-301-13	Sequence 13, Appli
38	717.5	50.1	248	1 US-07-901-707-5	Sequence 5, Appli
39	717.5	50.1	248	1 US-07-988-430-5	Sequence 5, Appli
40	717.5	50.1	248	1 US-08-425-336-5	Sequence 5, Appli
41	717.5	50.1	248	1 US-08-378-761A-75	Sequence 75, Appli
42	717.5	50.1	248	1 US-08-485-286-75	Sequence 75, Appli
43	717.5	50.1	248	1 US-08-488-113B-5	Sequence 5, Appli
44	717.5	50.1	248	1 US-08-477-484B-5	Sequence 5, Appli
45	717.5	50.1	248	2 US-08-646-360-5	Sequence 5, Appli

#### ALIGNMENTS

RESULT 1  
US-07-923-692C-4  
; Sequence 4, Application US/07923692C  
; Patent No. 5316931  
; GENERAL INFORMATION:  
; APPLICANT: Donson, Jon  
; APPLICANT: Dawson, William O.  
; APPLICANT: Grantham, George L.  
; APPLICANT: Turpen, Thomas H.  
; APPLICANT: Turpen, Ann Myers  
; APPLICANT: Garger, Stephen J.  
; APPLICANT: Grill, Laurence K.  
; TITLE OF INVENTION: RECOMBINANT PLANT VIRAL NUCLEIC ACIDS  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Limbach & Limbach  
; STREET: 2001 Ferry Building  
; CITY: San Francisco  
; STATE: CAL  
; ZIP: 94111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/923,692C  
; FILING DATE: 31-JUL-1992  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 600,244  
; FILING DATE: 22-OCT-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 641,617  
; FILING DATE: 16-JAN-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 310,881  
; FILING DATE: 17-FEB-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 160,766  
; FILING DATE: 26-FEB-1988  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 160,771  
; FILING DATE: 26-FEB-1988  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 347,637  
; FILING DATE: 05-MAY-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 363,138  
; FILING DATE: 08-JUN-1989

```
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 219,279
;; FILING DATE: 15-JUL-1988
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Halluin, Albert P.
;; REGISTRATION NUMBER: 28,957
;; REFERENCE/DOCKET NUMBER: BIOG-20121
;; REFERENCE/DOCKET NUMBER: USA
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 415-433-4150
;; TELEFAX: 415-433-8716
;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 289 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-07-923-692C-4

Query Match 98.7%; Score 1413; DB 1; Length 289;
Best Local Similarity 98.6%; Pred. No. 3.2e-137;
Matches 285; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MIRFLVLSLLILTLFTTPAVEGDVSPRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60
DB 1 MIRFLVLSLLILTLFTTPAVEGDVSPRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60
QY 61 RSSLPGSQRYALIHLTNYADETISVAIDVTNVIYMGYRAGDTSYFFNGASATEAAKYVFK 120
DB 61 RSSLPGSQRYALIHLTNYADETISVAIDVTNVIYMGYRAGDTSYFFNEASATEAAKYVFK 120
QY 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMVLIQS 180
DB 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMVLIQS 180
QY 181 TSEAAKYKFIQOIGKRVDTKFLPSLAIIISLNSWSALSQKIQIASTNNGQFESPVLIN 240
DB 181 TSEAAKYKFIQOIGKRVDTKFLPSLAIIISLNSWSALSQKIQIASTNNGQFETPVLIN 240
QY 241 AQNRVMTITNDAGVVTNSIALLLNRNNMAAMDDVPMTQSFSGCGSYAL 289
DB 241 AQNRVMTITNDAGVVTNSIALLLNRNNMAAMDDVPMTQSFSGCGSYAI 289

RESULT 2
US-08-184-237-4
; Sequence 4, Application US/08184237
; Patent No. 5589367
; GENERAL INFORMATION:
; APPLICANT: Donson, Jon
; APPLICANT: Dawson, William O.
; APPLICANT: Grantham, George L.
; APPLICANT: Turpen, Thomas H.
; APPLICANT: Turpen, Ann Myers
; APPLICANT: Garger, Stephen J.
; APPLICANT: Grill, Laurence K.
; TITLE OF INVENTION: RECOMBINANT PLANT VIRAL NUCLEIC ACIDS
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Limbach & Limbach
; STREET: 2001 Ferry Building
; CITY: San Francisco
; STATE: CAL
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/184,237
; FILING DATE:
; CLASSIFICATION: 435
```

APPLICANT: Dawson, William O.  
APPLICANT: Grantham, George L.  
APPLICANT: Turpen, Thomas H.  
APPLICANT: Turpen, Ann Myers  
APPLICANT: Garger, Stephen J.  
APPLICANT: Grill, Laurence K.  
TITLE OF INVENTION: RECOMBINANT PLANT VIRAL NUCLEIC ACIDS  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESS: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
FILING DATE: 07-JUNE-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 184,237  
FILING DATE: 19-JAN-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 600,244  
FILING DATE: 22-OCT-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 641,617  
FILING DATE: 16-JAN-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 310,881  
FILING DATE: 17-FEB-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 160,766  
FILING DATE: 26-FEB-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 160,771  
FILING DATE: 26-FEB-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 347,637  
FILING DATE: 05-MAY-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 363,138  
FILING DATE: 08-JUN-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 219,279  
FILING DATE: 15-JUL-1988  
ATTORNEY/AGENT INFORMATION:  
NAME: Halluin, Albert P.  
REGISTRATION NUMBER: 28,957  
REFERENCE/DOCKET NUMBER: 8129-112  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-854-3660  
TELEFAX: 415-854-3694  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 289 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-482-920-4

Query Match 98.7%; Score 1413; DB 2; Length 289;  
Best Local Similarity 98.6%; Pred. No. 3.2e-137;  
Matches 285; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 1 MIRFLVLSLLILTLTPPAVEGDVSFRLSGATSSSYGVFISNLKALPNERKLYDIPLL 60  
DB 1 MIRFLVLSLLILTLTPPAVEGDVSFRLSGATSSSYGVFISNLKALPNERKLYDIPLL 60  
QY 61 RSSLPQSQRVALIHLTNYADETISVAIDVTNVIYIMGRAGDTSYFFNGASATEAAKYVFK 120

Db 61 RSSLPQSQRVALIHLTNYADETISVAIDVTNVIYIMGRAGDTSYFFNGASATEAAKYVFK 120  
QY 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFFYNANGAASALMVLIOQS 180  
Db 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFFYNANGAASALMVLIOQS 180  
QY 181 TSEARYKFEIOQIGKRVDTFLPSLAIIISLNSWSALSQKIQIASTNNGQFETPVVLIN 240  
Db 181 TSEARYKFEIOQIGKRVDTFLPSLAIIISLNSWSALSQKIQIASTNNGQFETPVVLIN 240  
QY 241 AQNRVTITNDAGVVTNSIALLLNNNMAAMDDVPMTQSGCGSYAL 289  
Db 241 AQNRVTITNDAGVVTNSIALLLNNNMAAMDDVPMTQSGCGSYAI 289  
RESULT 4  
US-08-484-341-4  
Sequence 4, Application US/08484341  
GENERAL INFORMATION:  
APPLICANT: Donson, Jon  
Dawson, William O.  
Grantham, George L.  
Turpen, Thomas H.  
Turpen, Ann Myers  
Garger, Stephen J.  
Grill, Laurence K.  
TITLE OF INVENTION: RECOMBINANT PLANT VIRAL NUCLEIC ACIDS  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Limbach & Limbach  
STREET: 2001 Ferry Building  
CITY: San Francisco  
STATE: CAL  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,341  
FILING DATE: 07-Jun-1995  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/184,237  
FILING DATE: <Unknown>  
APPLICATION NUMBER: US 600,244  
FILING DATE: 22-OCT-1990  
APPLICATION NUMBER: US 641,617  
FILING DATE: 16-JAN-1991  
APPLICATION NUMBER: US 310,881  
FILING DATE: 17-FEB-1989  
APPLICATION NUMBER: US 160,766  
FILING DATE: 26-FEB-1988  
APPLICATION NUMBER: US 160,771  
FILING DATE: 26-FEB-1988  
APPLICATION NUMBER: US 347,637  
FILING DATE: 05-MAY-1989  
APPLICATION NUMBER: US 363,138  
FILING DATE: 08-JUN-1989  
APPLICATION NUMBER: US 219,279  
FILING DATE: 15-JUL-1988  
ATTORNEY/AGENT INFORMATION:  
NAME: Halluin, Albert P.  
REGISTRATION NUMBER: 28,957  
REFERENCE/DOCKET NUMBER: BIOG-20121 USA  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-433-4150  
TELEFAX: 415-433-8716  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 289 amino acids

```
;
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-08-484-341-4

Query Match      98.7%; Score 1413; DB 3; Length 289;
Best Local Similarity 98.6%; Pred. No. 3.2e-137;
Matches 285; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MIRELVLSLLILTLFTTPAVEGDSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60
DB 1 MIRELVLSLLILTLFTTPAVEGDSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60
QY 61 RSSLPGSORYALIHILTNVADETISVAIDVTNVYIMGYRAGDTSYFFNGASATEAAKYVK 120
DB 61 RSSLPGSORYALIHILTNVADETISVAIDVTNVYIMGYRAGDTSYFFNGASATEAAKYVK 120
QY 121 DAMRKVTLPYSGNVERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMWLIQS 180
DB 121 DAMRKVTLPYSGNVERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMWLIQS 180
QY 181 TSEARYKFIEQIGKRVDTFLPSLAISLNSWSALSQIQIASTNNGQFETPVVLIN 240
DB 181 TSEARYKFIEQIGKRVDTFLPSLAISLNSWSALSQIQIASTNNGQFETPVVLIN 240
QY 241 AQNRQVTITNDAGVVTNSNIALLLNRNNMAAMDDDDVPMTQSFSGCGSYAL 289
DB 241 AQNRQVTITNDAGVVTNSNIALLLNRNNMAAMDDDDVPMTQSFSGCGSYAI 289

RESULT 5
US-08-483-502-4
; Sequence 4, Application US/08483502
; Patent No. 6284492
; GENERAL INFORMATION:
; APPLICANT: Donson, Jon
; APPLICANT: Dawson, William O.
; APPLICANT: Grantham, George L.
; APPLICANT: Turpen, Thomas H.
; APPLICANT: Turpen, Ann Myers
; APPLICANT: Garger, Stephen J.
; APPLICANT: Grill, Laurence K.
; TITLE OF INVENTION: RECOMBINANT PLANT VIRAL NUCLEIC ACIDS
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Venable, Baetjer, Howard & Civiletti
; STREET: 1201 New York Avenue N.W., Suite 1000
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483,502
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/739,143
; FILING DATE:
; APPLICATION NUMBER: US 600,244
; FILING DATE: 22-OCT-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 641,617
; FILING DATE: 16-JAN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 310,881
; FILING DATE: 17-FEB-1989
; PRIOR APPLICATION DATA:

;
; APPLICATION NUMBER: US 160,766
; FILING DATE: 26-FEB-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 160,771
; FILING DATE: 26-FEB-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 347,637
; FILING DATE: 05-MAY-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 363,138
; FILING DATE: 08-JUN-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 219,279
; FILING DATE: 15-JUL-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Ihnen, Jeffrey L.
; REGISTRATION NUMBER: 28,957
; REFERENCE/DOCKET NUMBER: 18604-090574
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-962-4810
; TELEFAX: 202-962-8300
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 289 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-483-502-4

Query Match      98.7%; Score 1413; DB 3; Length 289;
Best Local Similarity 98.6%; Pred. No. 3.2e-137;
Matches 285; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MIRELVLSLLILTLFTTPAVEGDSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60
DB 1 MIRELVLSLLILTLFTTPAVEGDSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60
QY 61 RSSLPGSORYALIHILTNVADETISVAIDVTNVYIMGYRAGDTSYFFNGASATEAAKYVK 120
DB 61 RSSLPGSORYALIHILTNVADETISVAIDVTNVYIMGYRAGDTSYFFNGASATEAAKYVK 120
QY 121 DAMRKVTLPYSGNVERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMWLIQS 180
DB 121 DAMRKVTLPYSGNVERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMWLIQS 180
QY 181 TSEARYKFIEQIGKRVDTFLPSLAISLNSWSALSQIQIASTNNGQFETPVVLIN 240
DB 181 TSEARYKFIEQIGKRVDTFLPSLAISLNSWSALSQIQIASTNNGQFETPVVLIN 240
QY 241 AQNRQVTITNDAGVVTNSNIALLLNRNNMAAMDDDDVPMTQSFSGCGSYAL 289
DB 241 AQNRQVTITNDAGVVTNSNIALLLNRNNMAAMDDDDVPMTQSFSGCGSYAI 289

RESULT 6
US-09-726-651A-4
; Sequence 4, Application US/09726651A
; Patent No. 6448046
; GENERAL INFORMATION:
; APPLICANT: Donson, Jon
; APPLICANT: Dawson, William O.
; APPLICANT: Grantham, George L.
; APPLICANT: Turpen, Thomas H.
; APPLICANT: Turpen, Ann M.
; APPLICANT: Garger, Stephen J.
; APPLICANT: Grill, Laurence K.
; TITLE OF INVENTION: RECOMBINANT ANIMAL VIRUS NUCLEIC ACIDS
; FILE REFERENCE: 008010023CNUS01
; CURRENT APPLICATION NUMBER: US/09/726,651A
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: 08/483,502
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 08/184,237
```

;; PRIOR FILING DATE: 1994-01-19  
;; PRIOR APPLICATION NUMBER: 07/923,692  
;; PRIOR FILING DATE: 1992-07-31  
;; PRIOR APPLICATION NUMBER: 07/600,244  
;; PRIOR FILING DATE: 1990-10-22  
;; PRIOR APPLICATION NUMBER: 07/641,617  
;; PRIOR FILING DATE: 1991-01-16  
;; PRIOR APPLICATION NUMBER: 07/737,899  
;; PRIOR FILING DATE: 1991-07-26  
;; NUMBER OF SEQ ID NOS: 19  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO. 4  
;; LENGTH: 289  
;; TYPE: PRT  
;; ORGANISM: Chinese Cucumber alpha-trichosanthin  
US-09-726-651A-4

Query Match 98.7%; Score 1413; DB 4; Length 289;  
Best Local Similarity 98.6%; Pred. No. 3.2e-137;  
Matches 285; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 MRFVLVLSLLLTFTLTTPAVEGDVSRFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60  
Db 1 MRFVLVLSLLLTFTLTTPAVEGDVSRFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60  
  
QY 61 RSSLPGSQRYALIHLYNVADETISVAIDVTNVYIMGYRAGDTSYFFNGASATEAAKYVFK 120  
Db 61 RSSLPGSQRYALIHLYNVADETISVAIDVTNVYIMGYRAGDTSYFFNGASATEAAKYVFK 120  
  
QY 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMWLIQS 180  
Db 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMWLIQS 180  
  
QY 181 TSEARYKFIEQOIGKRVDTKFLPSLAIISLNSWSALSQIQIASTNNGQFESPVLIN 240  
Db 181 TSEARYKFIEQOIGKRVDTKFLPSLAIISLNSWSALSQIQIASTNNGQFESPVLIN 240  
  
QY 241 AQNRVTITNDAGVVTNSIALLLNRNNMAAMDDDDVPMTQSFSGCSYAL 289  
Db 241 AQNRVTITNDAGVVTNSIALLLNRNNMAAMDDDDVPMTQSFSGCSYAI 289

RESULT 7  
US-08-378-761A-74  
; Sequence 74, Application US/08378761A  
; Patent No. 5635384

;; GENERAL INFORMATION:  
;; APPLICANT: WALSH, TERENCE A  
;; APPLICANT: HEY, TIMOTHY D  
;; APPLICANT: MORGAN, ALICE ER  
;; TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE  
;; TITLE OF INVENTION: PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF  
;; TITLE OF INVENTION: USING  
;; NUMBER OF SEQUENCES: 81  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: ANDREA T. BORUCKI  
;; STREET: 9330 ZIONSVILLE ROAD  
;; CITY: INDIANAPOLIS  
;; STATE: IN  
;; COUNTRY: US  
;; ZIP: 46268  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/378,761A  
;; FILING DATE: 26-JAN-1995  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: BORUCKI, ANDREA T  
;; REGISTRATION NUMBER: 33651

;; REFERENCE/DOCKET NUMBER: 38272B  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (317) 337-4846  
;; INFORMATION FOR SEQ ID NO: 74:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 267 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-08-378-761A-74

Query Match 92.2%; Score 1321; DB 1; Length 267;  
Best Local Similarity 99.3%; Pred. No. 8.4e-128;  
Matches 265; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 23 GDVSPRLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSSLPQSRYALIHLYNVADET 82  
Db 1 GDVSPRLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSSLPQSRYALIHLYNVADET 60  
  
QY 83 ISVAIDVTNVYIMGYRAGDTSYFFNGASATEAAKYVFKDAMRKVTLPYSGNYERLQTAAG 142  
Db 61 ISVAIDVTNVYIMGYRAGDTSYFFNGASATEAAKYVFKDAMRKVTLPYSGNYERLQTAAG 120  
  
QY 143 KIRENIPGLPALDSAITTLFYNNANSAASALMWLIQSTSEARYKFIEQOIGKRVDTKF 202  
Db 121 KIRENIPGLPALDSAITTLFYNNANSAASALMWLIQSTSEARYKFIEQOIGKRVDTKF 180  
  
QY 203 LPSLAIISLNSWSALSQIQIASTNNGQFESPVLINNAQNORVTITNDAGVVTNSIAL 262  
Db 181 LPSLAIISLNSWSALSQIQIASTNNGQFESPVLINNAQNORVTITNDAGVVTNSIAL 240  
  
QY 263 LLNRNNMAAMDDDDVPMTQSFSGCSYAL 289  
Db 241 LLNRNNMAAMDDDDVPMTQSFSGCSYAI 267

## RESULT 8

US-08-485-286-74  
; Sequence 74, Application US/08485286  
; Patent No. 5646026  
; Patent No. 5646026 5646119

;; GENERAL INFORMATION:  
;; APPLICANT: WALSH, TERENCE A  
;; APPLICANT: HEY, TIMOTHY D  
;; APPLICANT: MORGAN, ALICE ER  
;; TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE  
;; TITLE OF INVENTION: PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF  
;; TITLE OF INVENTION: USING  
;; NUMBER OF SEQUENCES: 81  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: ANDREA T. BORUCKI  
;; STREET: 9330 ZIONSVILLE ROAD  
;; CITY: INDIANAPOLIS  
;; STATE: IN  
;; COUNTRY: US  
;; ZIP: 46268  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/485,286  
;; FILING DATE:  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/378761  
;; FILING DATE: 26-JAN-1995  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: BORUCKI, ANDREA T  
;; REGISTRATION NUMBER: 33651  
;; REFERENCE/DOCKET NUMBER: 38272B

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (317) 337-4846  
INFORMATION FOR SEQ ID NO: 74:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 267 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-485-286-74

Query Match 92.2%; Score 1321; DB 1; Length 267;  
Best Local Similarity 99.3%; Pred. No. 8.4e-128;  
Matches 265; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 23 GDVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSLPGSQRYALIHLYNVADET 82  
DB 1 GDVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSLPGSQRYALIHLYNVADET 60

QY 83 ISVAIDVTNVYIMGYRAGDTSYFFNGASATEAAKVFVKDAMRKVTLTPYSGNYERLQTAAG 142  
DB 61 ISVAIDVTNVYIMGYRAGDTSYFFNEASATEAAKVFVKDAMRKVTLTPYSGNYERLQTAAG 120

QY 143 KIRENIPLGLPALDSAITTLFYNNANSAASALMWLIQSTSEAAKYKFIQQIGKRVDTKTF 202  
DB 121 KIRENIPLGLPALDSAITTLFYNNANSAASALMWLIQSTSEAAKYKFIQQIGKRVDTKTF 180

QY 203 LPSLAISLNSWSALSQIQIQTASTNNGQFSPVVLINAQNRVTITNVDAGVVTSNIAL 262  
DB 181 LPSLAISLNSWSALSQIQIQTASTNNGQFSPVVLINAQNRVTITNVDAGVVTSNIAL 240

QY 263 LLNRNNMAAMDDVPMTQSFSGGSAI 289  
DB 241 LLNRNNMAAMDDVPMTQSFSGGSAI 267

RESULT 9  
US-08-488-113B-6  
Sequence 6, Application US/08488113B  
Patent No. 5744580  
GENERAL INFORMATION:  
APPLICANT: Better, Marc D.  
APPLICANT: Carroll, Stephen F.  
APPLICANT: Studnika, Gary M.  
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
TITLE OF INVENTION: Proteins  
NUMBER OF SEQUENCES: 169  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
STREET: 500 West Madison Street, 34th floor  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60661  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/488,113B  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/425,336  
FILING DATE: 18-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/064,691  
FILING DATE: 12-MAY-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/988,430  
FILING DATE: 09-DEC-1992  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/787,567  
FILING DATE: 04-NOV-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: McNicholas, Janet M.  
REGISTRATION NUMBER: 32,918  
REFERENCE/DOCKET NUMBER: 11022US07/200-70.F3.C2A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/707-8889  
TELEFAX: 312/707-9155  
TELEX: 650 388-1248  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 247 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-488-113B-6

Query Match 83.9%; Score 1202; DB 1; Length 247;  
Best Local Similarity 98.8%; Pred. No. 1.3e-115;  
Matches 244; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 24 DVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSLPGSQRYALIHLYNVADETI 83  
DB 1 DVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSLPGSQRYALIHLYNVADETI 60

QY 84 SVADVTNVYIMGYRAGDTSYFFNGASATEAAKVFVKDAMRKVTLTPYSGNYERLQTAAG 143  
DB 61 SVADVTNVYIMGYRAGDTSYFFNEASATEAAKVFVKDAMRKVTLTPYSGNYERLQTAAG 120

QY 144 IRENIPGLPALDSAITTLFYNNANSAASALMWLIQSTSEAAKYKFIQQIGKRVDTKTF 203  
DB 121 IRENIPGLPALDSAITTLFYNNANSAASALMWLIQSTSEAAKYKFIQQIGKRVDTKTF 180

QY 204 PSLAIIISLNSWSALSQIQIQTASTNNGQFSPVVLINAQNRVTITNVDAGVVTSNIAL 263  
DB 181 PSLAIIISLNSWSALSQIQIQTASTNNGQFSPVVLINAQNRVTITNVDAGVVTSNIAL 240

QY 264 LLNRNNMA 270  
DB 241 LLNRNNMA 247

RESULT 10  
US-08-477-484B-6  
Sequence 6, Application US/08477484B  
Patent No. 5756699  
GENERAL INFORMATION:  
APPLICANT: Better, Marc D.  
APPLICANT: Carroll, Stephen F.  
APPLICANT: Studnika, Gary M.  
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
TITLE OF INVENTION: Proteins  
NUMBER OF SEQUENCES: 169  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
STREET: 500 West Madison Street, 34th floor  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60661  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/477,484B  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 530



```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/425,336
; FILING DATE: 18-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/064,691
; FILING DATE: 12-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/988,430
; FILING DATE: 09-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/901,707
; FILING DATE: 19-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/787,567
; FILING DATE: 04-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McNicholas, Janet M.
; REGISTRATION NUMBER: 32,918
; REFERENCE/DOCKET NUMBER: 11022US07/200-70.P3.C2A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/707-9155
; TELEFAX: 312/707-9155
; TELEX: 650 388-1248
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 247 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-477-484B-6

```

```

Query Match      83.9%; Score 1202; DB 1; Length 247;
Best Local Similarity 98.8%; Pred. No. 1.3e-115;
Matches 244; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 24 DVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSSLPFSQRYALIHLTNYADETI 83
Db 1 DVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSSLPFSQRYALIHLTNYADETI 60

QY 84 SVAIDVTNVIYMGYRAGDTSYFFNCGASATEAAKYVFKDAMRKVTLPYSGNYERLQTAACK 143
Db 61 SVAIDVTNVIYMGYRAGDTSYFFNCGASATEAAKYVFKDAMRKVTLPYSGNYERLQTAACK 120

QY 144 IRENIPLGLPALDSAITTLFYNNANSAASALMVLIOQSTSEAAKYKFIQQIGKRVDTKPL 203
Db 121 IRENIPLGLPALDSAITTLFYNNANSAASALMVLIOQSTSEAAKYKFIQQIGKRVDTKPL 180

QY 204 PSLAIISLNSWSALSQKIQIASTNNGQFSPVVLINAQNRVTTITNDAGVVTSNIAL 263
Db 181 PSLAIISLNSWSALSQKIQIASTNNGQFSPVVLINAQNRVTTITNDAGVVTSNIAL 240

QY 264 LNRNMA 270
Db 241 LNRNMA 247

```

```

RESULT 11
US-08-646-360-6
; Sequence 6, Application US/08646360
; Patent No. 5837491
; GENERAL INFORMATION:
; APPLICANT: Better, Marc D.
; APPLICANT: Carroll, Stephen F.
; APPLICANT: Studnika, Gary M.
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
; TITLE OF INVENTION: Proteins
; NUMBER OF SEQUENCES: 173
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.
; STREET: 500 West Madison Street, 34th floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA

```

```

; ZIP: 60661
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,360
; FILING DATE: 13-MAY-1996
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/05348
; FILING DATE: 12-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/064,691
; FILING DATE: 12-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/988,430
; FILING DATE: 09-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/901,707
; FILING DATE: 19-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/787,567
; FILING DATE: 04-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McNicholas, Janet M.
; REGISTRATION NUMBER: 32,918
; REFERENCE/DOCKET NUMBER: 200-70.P4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/707-8889
; TELEFAX: 312/707-9155
; TELEX: 650 388-1248
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 247 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-646-360-6

```

```

Query Match      83.9%; Score 1202; DB 2; Length 247;
Best Local Similarity 98.8%; Pred. No. 1.3e-115;
Matches 244; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 24 DVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSSLPFSQRYALIHLTNYADETI 83
Db 1 DVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSSLPFSQRYALIHLTNYADETI 60

QY 84 SVAIDVTNVIYMGYRAGDTSYFFNCGASATEAAKYVFKDAMRKVTLPYSGNYERLQTAACK 143
Db 61 SVAIDVTNVIYMGYRAGDTSYFFNCGASATEAAKYVFKDAMRKVTLPYSGNYERLQTAACK 120

QY 144 IRENIPLGLPALDSAITTLFYNNANSAASALMVLIOQSTSEAAKYKFIQQIGKRVDTKPL 203
Db 121 IRENIPLGLPALDSAITTLFYNNANSAASALMVLIOQSTSEAAKYKFIQQIGKRVDTKPL 180

QY 204 PSLAIISLNSWSALSQKIQIASTNNGQFSPVVLINAQNRVTTITNDAGVVTSNIAL 263
Db 181 PSLAIISLNSWSALSQKIQIASTNNGQFSPVVLINAQNRVTTITNDAGVVTSNIAL 240

QY 264 LNRNMA 270
Db 241 LNRNMA 247

```

```

RESULT 12
US-08-839-765-6
; Sequence 6, Application US/08839765
; Patent No. 6146631
; GENERAL INFORMATION:
; APPLICANT: Better, Marc D.
; APPLICANT: Carroll, Stephen F.

```

APPLICANT: Studnika, Gary M.  
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
TITLE OF INVENTION: Proteins  
NUMBER OF SEQUENCES: 169  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
STREET: 500 West Madison Street, 34th floor  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60661  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/839,765  
FILING DATE: 15-APR-1997  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/425,336  
FILING DATE: 18-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/064,691  
FILING DATE: 12-MAY-1993  
APPLICATION DATA:  
APPLICATION NUMBER: US 07/988,430  
FILING DATE: 09-DEC-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/787,567  
FILING DATE: 04-NOV-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: McNicholas, Janet M.  
REGISTRATION NUMBER: 32,918  
REFERENCE/DOCKET NUMBER: 11022US09/200-70.P3.C3  
TELEPHONE: 312/707-9155  
TELEFAX: 312/707-9155  
TELEX: 650 388-1248  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 247 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-839-765-6

Query Match 83.9%; Score 1202; DB 3; Length 247;  
Best Local Similarity 98.8%; Pred. No. 1.3e-115;  
Matches 244; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 24 DVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSLPGSQRYALIHLTNYADETI 83  
DB 1 DVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSLPGSQRYALIHLTNYADETI 60

QY 84 SVAIDVTNYIMGYRAGDTSYFPNGASATEAAKYVFKDAMRKVTLTPYSGNYERLQTAAGK 143  
DB 61 SVAIDVTNYIMGYRAGDTSYFPNGASATEAAKYVFKDAMRKVTLTPYSGNYERLQTAAGK 120

QY 144 IRENIPILGUPALDSAITTLFYNNANSAALMWLIQSTSEAAKYKTEQQIGKRVKDTFL 203  
DB 121 IRENIPILGUPALDSAITTLFYNNANSAALMWLIQSTSEAAKYKTEQQIGKRVKDTFL 180

QY 204 PSIAISLNSWALSQKIQIASTNNGQFSPVVLINAQNRQVTIINVDAGVVTNSIAL 263  
DB 181 PSIAISLNSWALSQKIQIASTNNGQFSPVVLINAQNRQVTIINVDAGVVTNSIAL 240

QY 264 LNRNMA 270  
DB 1 LNRNMA 270

Db 241 LNRNMA 247

RESULT 13  
US-09-136-389-6  
Sequence 6, Application US/09136389  
Patent No. 6146850  
GENERAL INFORMATION:  
APPLICANT: Better, Marc D.  
APPLICANT: Carroll, Stephen F.  
APPLICANT: Studnika, Gary M.  
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
NUMBER OF SEQUENCES: 173  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
STREET: 500 West Madison Street, 34th floor  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60661  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/136,389  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/646,360  
FILING DATE: 13-MAY-1996  
APPLICATION NUMBER: PCT/US94/05348  
FILING DATE: 12-MAY-1994  
APPLICATION DATA:  
APPLICATION NUMBER: US 08/064,691  
FILING DATE: 12-MAY-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/988,430  
FILING DATE: 09-DEC-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/787,567  
FILING DATE: 04-NOV-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: McNicholas, Janet M.  
REGISTRATION NUMBER: 32,918  
REFERENCE/DOCKET NUMBER: 200-70.P4  
TELEPHONE: 312/707-8889  
TELEFAX: 312/707-9155  
TELEX: 650 388-1248  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 247 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-136-389-6

Query Match 83.9%; Score 1202; DB 3; Length 247;  
Best Local Similarity 98.8%; Pred. No. 1.3e-115;  
Matches 244; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 24 DVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSLPGSQRYALIHLTNYADETI 83  
DB 1 DVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSLPGSQRYALIHLTNYADETI 60

QY 84 SVAIDVTNYIMGYRAGDTSYFPNGASATEAAKYVFKDAMRKVTLTPYSGNYERLQTAAGK 143  
DB 1 LNRNMA 270

```

Db 61 SVAIDVTNVIYMGYRAGDTSYFFNEASATEAAKYVFDAMRKVTLPYSGNYERLQTAAGK 120
QY 144 IRENIPGLPALDSAITTLFFYNANSAASALMWLIQSTSEAAKYKFIQQIGKRVDTKFL 203
Db 121 IRENIPGLPALDSAITTLFFYNANSAASALMWLIQSTSEAAKYKFIQQIGKRVDTKFL 180
QY 204 PSIAIISLENSWSALSQIOIASTNNQGFSPVVLINAQNRVTITNVDAGVVTSNIAL 263
Db 181 PSIAIISLENSWSALSQIOIASTNNQGFSPVVLINAQNRVTITNVDAGVVTSNIAL 240
QY 264 LNRNNMA 270
Db 241 LNRNNMA 247

RESULT 14
US-09-610-838-6
; Sequence 6, Application US/09610838
; Patent No. 6376217
; GENERAL INFORMATION:
; APPLICANT: Better, Marc D.
; APPLICANT: Carroll, Stephen F.
; APPLICANT: Studnika, Gary M.
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
; TITLE OF INVENTION: Proteins
; NUMBER OF SEQUENCES: 173
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.
; STREET: 500 West Madison Street, 34th floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60661
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/610,838
; FILING DATE: 06-JUL-2000
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/136,389
; FILING DATE: 18-AUG-1998
; APPLICATION NUMBER: 08/646,360
; FILING DATE: 13-MAY-1996
; APPLICATION NUMBER: PCT/US94/05348
; FILING DATE: 12-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/064,691
; FILING DATE: 12-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/988,430
; FILING DATE: 09-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/901,707
; FILING DATE: 19-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/787,567
; FILING DATE: 04-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McNicholas, Janet M.
; REGISTRATION NUMBER: 32,918
; REFERENCE/DOCKET NUMBER: 200-70.P4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/707-8889
; TELEFAX: 312/707-9155
; TELEX: 650 388-1248
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 247 amino acids
; TYPE: amino acid

```

```

; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-610-838-6
Query Match 83.9%; Score 1202; DB 4; Length 247;
Best Local Similarity 98.8%; Pred. No. 1.3e-115;
Matches 244; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 24 DVSFRLSGATSSSYGVFISNLKALPNERKLYDIPILRSLPGSQRYALHLTNVADETI 83
Db 1 DVSFRLSGATSSSYGVFISNLKALPNERKLYDIPILRSLPGSQRYALHLTNVADETI 60
QY 84 SVAIDVTNVIYMGYRAGDTSYFFNEASATEAAKYVFDAMRKVTLPYSGNYERLQTAAGK 143
Db 61 SVAIDVTNVIYMGYRAGDTSYFFNEASATEAAKYVFDAMRKVTLPYSGNYERLQTAAGK 120
QY 144 IRENIPGLPALDSAITTLFFYNANSAASALMWLIQSTSEAAKYKFIQQIGKRVDTKFL 203
Db 121 IRENIPGLPALDSAITTLFFYNANSAASALMWLIQSTSEAAKYKFIQQIGKRVDTKFL 180
QY 204 PSIAIISLENSWSALSQIOIASTNNQGFSPVVLINAQNRVTITNVDAGVVTSNIAL 263
Db 181 PSIAIISLENSWSALSQIOIASTNNQGFSPVVLINAQNRVTITNVDAGVVTSNIAL 240
QY 264 LNRNNMA 270
Db 241 LNRNNMA 247

RESULT 15
US-09-711-485-6
; Sequence 6, Application US/09711485
; Patent No. 6649742
; GENERAL INFORMATION:
; APPLICANT: Better, Marc D.
; APPLICANT: Carroll, Stephen F.
; APPLICANT: Studnika, Gary M.
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
; TITLE OF INVENTION: Proteins
; NUMBER OF SEQUENCES: 169
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.
; STREET: 500 West Madison Street, 34th floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60661
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/711,485
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/839,765
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/064,691
; FILING DATE: 12-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/988,430
; FILING DATE: 09-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/901,707
; FILING DATE: 19-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/787,567
; FILING DATE: 04-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McNicholas, Janet M.

```

Wed Sep 15 10:32:03 2004

REGISTRATION NUMBER: 32,918  
REFERENCE/DOCKET NUMBER: 11022US09/200-70.P3.C3

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312/707-8889  
TELEFAX: 312/707-9155  
TELEX: 650 388-1248

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 247 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-711-485-6

Query Match 83.9%; Score 1202; DB 4; Length 247;

Best Local Similarity 98.8%; Pred. No. 1.3e-115;

Matches 244; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 24 DVSRFLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSSLPQSORYALIHLTNYADETI 83

Db 1 DVSRFLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSSLPQSORYALIHLTNYADETI 60

QY 84 SVAIDVTNVIYIMGYRAGDTSYFFNGASATEAAKYVFKDAMRKVTLTPYSGNYERLQTAAGK 143

Db 61 SVAIDVTNVIYIMGYRAGDTSYFFNEASATEAAKYVFKDAMRKVTLTPYSGNYERLQTAAGK 120

QY 144 IRENIPGLPALDSAITTLFYNNANSAASALMVLIOQSTSEARYKFEIQQIGKRVDTKFL 203

Db 121 IRENIPGLPALDSAITTLFYNNANSAASALMVLIOQSTSEARYKFEIQQIGKRVDTKFL 180

QY 204 PSIAIISLENSWSALSQKIQIASTNNGQFESPVVLINAQNRVTITNVDAGVVTSGNIALL 263

Db 181 PSIAIISLENSWSALSQKIQIASTNNGQFESPVVLINAQNRVTITNVDAGVVTSGNIALL 240

QY 264 LNRNNMA 270

Db 241 LNRNNMA 247

Search completed: September 10, 2004, 14:02:34

Job time : 35 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 13:24:40 ; Search time 121 Seconds  
(without alignments)  
674.845 Million cell updates/sec

Title: US-09-905-247-1

Perfect score: 1432

Sequence: 1 MRFLVLSLLILFLTTTPA.....AAMDDVPMTQSGCGSYAL 289

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: Genesep29Jan04.\*
- 2: Genesep1980s.\*
- 3: Genesep1990s.\*
- 4: Genesep2000s.\*
- 5: Genesep2001s.\*
- 6: Genesep2002s.\*
- 7: Genesep2003as.\*
- 8: Genesep2003bs.\*
- 9: Genesep2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1424	99.4	289	5	ABB07660 Native tr
2	1422	99.3	289	4	AAB99329 Trichosan
3	1413	98.7	289	2	AAR07514 Trichosan
4	1413	98.7	289	2	AAR25572 Trichosan
5	1413	98.7	289	2	AAR29272 Trichosan
6	1413	98.7	289	2	AAR55129 Trichosan
7	1413	98.7	289	2	AAR10468 Trichosan
8	1413	98.7	289	2	AAR11870 Trichosan
9	1413	98.7	289	2	AAR01374 Trichosan
10	1413	98.7	289	2	AAR84192 Trichosan
11	1413	98.7	289	3	AAR87791 Trichosan
12	1410	98.5	289	2	AAR32986 Trichosan
13	1326	92.6	289	2	AAR07523 Trichosan
14	1326	92.6	289	2	AAR29276 Trichosan
15	1321	92.2	267	2	AAR21703 Trichosan
16	1321	92.2	267	2	AAR25140 Trichosan
17	1214	84.8	247	3	AAR69048 Trichosan
18	1214	84.8	248	2	AAR07518 Trichosan
19	1214	84.8	248	2	AAR25573 Trichosan
20	1211	84.6	247	2	AAR67359 Trichosan
21	1183	82.6	247	2	AAR21605 Trichosan
22	1164.5	81.3	290	2	AAR92481 Trichosan
23	1144	79.9	255	2	AAR74181 Trichosan
24	1139	79.5	246	2	AAR52636 Trichosan
25	1130	78.9	255	2	AAR37295 Trichosan

26	1130	78.9	255	2	AAR63907 Type I ri
27	1089.5	76.1	287	3	AAB01299 Wild type
28	1074	75.0	496	4	AAB36828 BDI-G28.5
29	1064	74.3	247	3	AAY69043 Amino aci
30	1064	74.3	248	4	AAB36824 Residues
31	1062.5	74.2	267	3	AAB01300 Altered b
32	806.5	56.3	263	2	AAR63908 Type I ri
33	806.5	56.3	263	2	AAR74182 Type I ri
34	802.5	56.0	277	2	AAR29909 Prod. of
35	799.5	55.8	283	2	AAR37296 Plant typ
36	798	55.7	280	2	AAR07520 Alpha-Tri
37	798	55.7	280	2	AAR07521 Alpha-Tri
38	798	55.7	280	2	AAR25575 Ribosome
39	798	55.7	280	2	AAR25576 Ribosome
40	786.5	54.9	272	2	AAR07522 Alpha-Tri
41	786.5	54.9	272	2	AAR25577 Ribosome
42	775	54.1	278	2	AAR29910 Prod. of
43	772.5	53.9	276	2	AAR53731 Luffin (r
44	760	53.1	250	2	AAR21707 Momordin.
45	760	53.1	250	2	AAR25144 Momordin

## ALIGNMENTS

RESULT 1  
ABB07660  
ID ABB07660 standard; protein; 289 AA.  
XX  
AC ABB07660;  
XX  
DT 20-MAY-2002 (first entry)  
XX  
DE Native trichosanthin (TCS) protein sequence.  
XX  
KW Trichosanthin; TCS; mutant; MTCS; therapeutic; cytostatic; antitumour;  
KW anti-human immunodeficiency virus; virucide; immunostimulant; gene;  
KW ectopic pregnancy.  
XX  
OS Trichosanthin kirilowii.

Key	Location/Qualifiers
Peptide	1..23
Misc-difference	/note= "signal peptide"
Protein	/note= "encoded by GAC"
Region	174..180
Misc-difference	/note= "mature protein"
Region	/note= "MTCS contains a modification of at least one amino acid residue in this region"
Misc-difference	175
Region	/note= "encoded by ATT"
Region	203..226
Misc-difference	/note= "MTCS contains a modification of at least one amino acid residue in this region"
Region	230..244
Misc-difference	/note= "MTCS contains a modification of at least one amino acid residue in this region"
Region	/note= "encode dby ATT"

W0200212537-A2.

14-FEB-2002.

18-JUL-2001; 2001WO-CN001178.

02-AUG-2000; 2000CN-00119553.

18-JAN-2001; 2001CN-00103102.

(BEIJ-) BEIJING STM BIOTECH LTD.  
(KEYJ/) KE Y.

```
PA (NIEH/) NIE H.
XX
XX PI Ke Y, Nie H;
XX
XX PF WPI; 2002-227165/28.
XX
XX DR N-PSDB; ABA95171.
XX
XX PA Mutant trichosanthin protein of low antigenicity useful for treating
XX
XX PT tumor e.g. leukemia, comprises a sequence of native trichosanthin with
XX
XX PT modification of at least one amino acid residue in three specific
XX
XX PT regions.
XX
XX XX Example 1; Fig 1; 42pp; English.
XX
XX PS The invention relates to a mutant trichosanthin (MTCs) protein of low
XX
XX CC antigenicity comprising a sequence of native TCS with the modification of
XX
XX CC at least one amino acid residue in three regions which is 174-180, 203-
XX
XX CC 227 and 230-244 and substantially retaining the biological activities of
XX
XX CC the native TCS. The MTCs protein, its fragment or derivative is useful as
XX
XX CC an therapeutic agent; or for preparing a medicament for treating viral
XX
XX CC disease e.g. acquired immunodeficiency syndrome (AIDS), tumour e.g.
XX
XX CC leukemia, for inducing abortion and/or treating ectopic pregnancy. The
XX
XX CC present sequence represents the native TCS
XX
XX SQ Sequence 289 AA;
XX
XX Query Match 99.4%; Score 1424; DB 5; Length 289;
XX
XX Best Local Similarity 99.7%; Pred. No. 1.7e-127;
XX
XX Matches 288; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 MIRELVLSLLILTLFTTPAVEGDSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60
XX
XX DB 1 MIRELVLSLLILTLFTTPAVEGDSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60
XX
XX QY 61 RSSLPGSORYALIHLTNYADETISVAIDVTNVYINGYRAGDTSYFFNGASATEAAKYVFK 120
XX
XX DB 61 RSSLPGSORYALIHLTNYADETISVAIDVTNVYINGYRAGDTSYFFNGASATEAAKYVFK 120
XX
XX QY 121 DAMRKVTLPYSGNYERLQTAAGKTRENTPLGLPALDSAITTLFYNNANSAASALMVLIQS 180
XX
XX DB 121 DAMRKVTLPYSGNYERLQTAAGKTRENTPLGLPALDSAITTLFYNNANSAASALMVLIQS 180
XX
XX QY 181 TSEAAARYKFEIQIGKRVDTKFLPSLAISLNSWSALSQIQIASTNNGQFESPVLIN 240
XX
XX DB 181 TSEAAARYKFEIQIGKRVDTKFLPSLAISLNSWSALSQIQIASTNNGQFESPVLIN 240
XX
XX QY 241 AQNRQVTITNDAGVVTNSNIALLLNRRNNMAAMDDVPMTQSFSGCGSYAL 289
XX
XX DB 241 AQNRQVTITNDAGVVTNSNIALLLNRRNNMAAMDDVPMTQSFSGCGSYAL 289
XX
XX RESULT 2
XX
XX ID AAB99329 standard; protein; 289 AA.
XX
XX AC AAB99329;
XX
XX XX 23-AUG-2001 (first entry)
XX
XX DT Trichosanthes kirilowii trichosanthin (TCS) protein sequence.
XX
XX DE Trichosanthes kirilowii; trichosanthin; TCS; mutagenesis; mutation; MTCs;
XX
XX KW mutant of trichosanthin; Mongolian snake-gourd; bioactivity; selectivity;
XX
XX KW cancer; virus; HIV; metaphase induced labour.
XX
XX XX Trichosanthes kirilowii.
XX
XX OS Trichosanthes kirilowii.
XX
XX XX Key Location/Qualifiers
XX
XX FH Misc-difference 230
XX
XX FT /note= "encoded by GCA"
XX
XX FT CN1283630-A.
XX
XX XX
```

```
PD 14-FEB-2001.
XX
XX PF 02-AUG-2000; 2000CN-00119553.
XX
XX PR 02-AUG-2000; 2000CN-00119553.
XX
XX XX (SHAN-) SHANGHAI INST CYTOBIOLOGY CHINESE ACAD.
XX
XX PI Ke Y, Nie H;
XX
XX XX WPI; 2001-291745/31.
XX
XX DR N-PSDB; AAH41473.
XX
XX XX Trichosanthin mutant and its preparing process.
XX
XX PT Disclosure; Page 2 (disclosure); 15pp; Chinese.
XX
XX PS The present invention describes a trichosanthin mutant which is prepared
XX
XX CC through the mutational deformation of the trichosanthin gene and using an
XX
XX CC expression system. The trichosanthin gene is isolated from Trichosanthes
XX
XX CC kirilowii (Mongolian snake-gourd). The trichosanthin mutant has several
XX
XX CC advantages including high bioactivity, high selectivity to target, and
XX
XX CC strong kill action to cancer cells, virus and HIV. It can also be used
XX
XX CC for metaphase induced labour. The present sequence represents the protein
XX
XX CC sequence of wild type trichosanthin which is given in the exemplification
XX
XX CC of the present invention
XX
XX SQ Sequence 289 AA;
XX
XX Query Match 99.3%; Score 1422; DB 4; Length 289;
XX
XX Best Local Similarity 99.3%; Pred. No. 2.6e-127;
XX
XX Matches 287; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 MIRELVLSLLILTLFTTPAVEGDSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60
XX
XX DB 1 MIRELVLSLLILTLFTTPAVEGDSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60
XX
XX QY 61 RSSLPGSORYALIHLTNYADETISVAIDVTNVYINGYRAGDTSYFFNGASATEAAKYVFK 120
XX
XX DB 61 RSSLPGSORYALIHLTNYADETISVAIDVTNVYINGYRAGDTSYFFNGASATEAAKYVFK 120
XX
XX QY 121 DAMRKVTLPYSGNYERLQTAAGKTRENTPLGLPALDSAITTLFYNNANSAASALMVLIQS 180
XX
XX DB 121 DAMRKVTLPYSGNYERLQTAAGKTRENTPLGLPALDSAITTLFYNNANSAASALMVLIQS 180
XX
XX QY 181 TSEAAARYKFEIQIGKRVDTKFLPSLAISLNSWSALSQIQIASTNNGQFESPVLIN 240
XX
XX DB 181 TSEAAARYKFEIQIGKRVDTKFLPSLAISLNSWSALSQIQIASTNNGQFESPVLIN 240
XX
XX QY 241 AQNRQVTITNDAGVVTNSNIALLLNRRNNMAAMDDVPMTQSFSGCGSYAL 289
XX
XX DB 241 AQNRQVTITNDAGVVTNSNIALLLNRRNNMAAMDDVPMTQSFSGCGSYAL 289
XX
XX RESULT 3
XX
XX ID AAR07514 standard; protein; 289 AA.
XX
XX AC AAR07514;
XX
XX XX 06-FEB-1991 (first entry)
XX
XX DT Trichosanthin from Trichosanthes kirilowii.
XX
XX DE trichosanthin (TCS); Human Immunodeficiency Virus; inhibition.
XX
XX KW Trichosanthes kirilowii.
XX
XX OS Trichosanthes kirilowii.
XX
XX XX Key Location/Qualifiers
XX
XX FH 1. .21
XX
XX FT /label= "signal peptide"
XX
XX FT /note= "hydrophobic"
XX
XX XX
```

PN WO9012097-A.  
 XX 18-OCT-1990.  
 XX 04-APR-1989; 89US-00333184.  
 PF 04-APR-1989; 89US-00333184.  
 PR 04-APR-1989; 89US-00333184.  
 XX (GENE-) GENELABS INC.  
 PA Piatek M, Chow T, Fry K;  
 XX WPI; 1990-334847/44.  
 DR N-PSDB; AAQ06343.  
 XX Recombinant tri:chosanthin protein - with selective inhibitory effect on  
 FT viral expression in HIV infected T-cells or monocyte-macrophase.  
 XX Example; Fig 4; 102pp; English.  
 CC Genomic DNA was isolated from T.kirilowii leaves from Korea and a library  
 CC was constructed. Clone p021D was identified as likely to contain a TCS-  
 CC encoding sequence in its 4kb insert. The deduced amino acid sequence is  
 CC identical to that of TCS purified from Cantonese T.kirilowii roots,  
 CC except for 2 conservative substitutions, i.e. Thr for Ser at position  
 CC 211 and Met for Thr at position 224. The Canton protein lacks the last 19  
 CC C-terminal amino acid residues. See also AAQ06344-Q06351  
 XX Sequence 289 AA;  
 SQ

Query Match 98.7%; Score 1413; DB 2; Length 289;  
 Best Local Similarity 98.6%; Pred. No. 1.9e-126;  
 Matches 285; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MIRFLVLSLLILTLTPAVEGDVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60  
 Db 1 MIRFLVLSLLILTLTPAVEGDVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60  
 QY 61 RSLPGSQRYALIHNTVADETISVAIDVTNVIYIMGYRAGDTSYFFNGASATEAAKYVFK 120  
 Db 61 RSLPGSQRYALIHNTVADETISVAIDVTNVIYIMGYRAGDTSYFFNGASATEAAKYVFK 120  
 QY 121 DAMRKVTLPYSGNYERLQTAAGKIRENTPLGLPALDSAITTLFYNNANSAALMVLIIQS 180  
 Db 121 DAMRKVTLPYSGNYERLQTAAGKIRENTPLGLPALDSAITTLFYNNANSAALMVLIIQS 180  
 QY 181 TSEAAKYKFIQQIGKRVDTKFLPSLAIISLNSWSALSKOIQTASTNNGOFESPVLIN 240  
 Db 181 TSEAAKYKFIQQIGKRVDTKFLPSLAIISLNSWSALSKOIQTASTNNGOFESPVLIN 240  
 QY 241 AQORVITITNDAGVITSNIALLNRRNNMAAMDDVPMTQSGGGSYAL 289  
 Db 241 AQORVITITNDAGVITSNIALLNRRNNMAAMDDVPMTQSGGGSYAI 289

RESULT 4  
 AAR25572  
 ID AAR25572 standard; protein; 289 AA.  
 XX  
 AC AAR25572;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 13-JAN-1993 (first entry)  
 XX  
 XX Trichosanthin from Trichosanthin kirilowii.  
 DE TCS; alpha-trichosanthin; Radix trichosanthin; abortifacient;  
 XX ribosome inactivating protein; RIP; HIV-infected human T cells;  
 KW human immunodeficiency virus.  
 XX Trichosanthin kirilowii.  
 OS  
 XX Key Location/Qualifiers  
 FH

Misc-difference 57 /note= "Leu in previously published sequence"  
 FT  
 FT Misc-difference 60 /note= "Ile in previously published sequence"  
 FT  
 FT Misc-difference 72 /note= "Ile in previously published sequence"  
 FT  
 FT Misc-difference 82..84 /note= "Ile in previously published sequence"  
 FT  
 FT /note= "region not present in previously published  
 FT sequence"  
 FT Misc-difference 92..93 /note= "previously published sequence contained a 10  
 FT amino acid insert (DAGLPNRAVL) between Val and Tyr"  
 FT  
 FT Misc-difference 143 /note= "Gly in previously published sequence"  
 FT  
 FT Misc-difference 144 /note= "Leu in previously published sequence"  
 FT  
 FT Misc-difference 196 /note= "Ser in previously published sequence"  
 FT  
 FT Misc-difference 214..215 /note= "previously published sequence contained a Leu  
 FT inserted between Ser and Trp"  
 FT  
 FT Misc-difference 216 /note= "Leu in previously published sequence"  
 FT  
 FT Misc-difference 231 /note= "Thr in previously published sequence"  
 FT  
 FT Misc-difference 234 /note= "Ser in directly sequenced TCS"  
 FT  
 FT Misc-difference 246..266 /note= "21 amino acids not present in previously  
 FT published sequence"  
 FT  
 FT Misc-difference 247 /note= "Thr in directly sequenced TCS"  
 FT  
 XX US5128460-A.  
 XX  
 XX 07-JUL-1992.  
 XX  
 PF 04-APR-1990; 90US-00504775.  
 XX  
 PR 04-APR-1989; 89US-00333184.  
 PR 07-SEP-1989; 89US-00404326.  
 XX (GENE-) GENELABS INC.  
 XX  
 XX Piatek M, Chow TP, Fry K;  
 XX WPI; 1992-249485/30.  
 DR N-PSDB; AAQ26499.  
 XX  
 XX Nucleic acid encoding trichosanthin protein - which can be used to  
 PT inactivate ribosome(s) to inhibit protein synthesis or to inhibit HIV  
 PT expression.  
 XX  
 XX Claim 1; Fig 4; 53pp; English.  
 XX  
 XX Trichosanthin protein was isolated from T.kirilowii root tuber and  
 CC sequenced. The amino acid sequence was used to design sets of degenerate  
 CC primers (see AAQ26506-8) for PCR amplification of the TCS coding  
 CC sequence. The amplified product was used as a probe to isolate TCS coding  
 CC sequence from T.kirilowii genomic libraries. One clone (p021D) contained  
 CC a 4kb insert. The amino acid sequence deduced from the p021D insert  
 CC differed from the purified TCS sequence by conservative substitutions at  
 CC two positions (see Features Table). The differences are postulated to  
 CC indicate minor variations between strains; the purified TCS was obtained  
 CC from the Canton region of China and the genomic DNA was obtained from  
 CC T.kirilowii leaves from Korea. The amino acid sequence was also found to  
 CC differ substantially from the previously published TCS sequence (Acta  
 CC Chemica Sinica, 43:1943, 1984 and Pure and Appl. Chem., 58(5):789, 1986).  
 CC The differences are shown in the Features Table. The present sequence  
 CC agrees closely with X-ray diffraction data on crystallized TCS. The 21-  
 CC amino acid insert also provides greater sequence homology with a number  
 CC of RIPS such as ricin A chain and abrin A chain than the previously  
 CC published sequence. (Updated on 25-MAR-2003 to correct PF field.)

CC (Updated on 25-MAR-2003 to correct PR field.)  
 XX Sequence 289 AA;  
 SQ

Query Match 98.7%; Score 1413; DB 2; Length 289;  
 Best Local Similarity 98.6%; Pred. No. 1.9e-126; Indels 0; Gaps 0;  
 Matches 285; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MIRFLVLSLLILFLTPPAVEGDVSRFGATSSSYGVFISNLRKALPNERKLYDIPLL 60  
 Db 1 MIRFLVLSLLILFLTPPAVEGDVSRFGATSSSYGVFISNLRKALPNERKLYDIPLL 60

Qy 61 RSSLPQSGQRYALIHNTYADETISVAIDVTNVYIMGYRAGDTSYFFNGASATEAAKYVFK 120  
 Db 61 RSSLPQSGQRYALIHNTYADETISVAIDVTNVYIMGYRAGDTSYFFNEASATEAAKYVFK 120

Qy 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPLGLPALDSAITTLFYNNANSASALMWLIQS 180  
 Db 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPLGLPALDSAITTLFYNNANSASALMWLIQS 180

Qy 181 TSEAAARYKFIEQQIGKRVDTFLPSLAIISLENSWSALSQIOIQTASTNNGQFESPVLIN 240  
 Db 181 TSEAAARYKFIEQQIGKRVDTFLPSLAIISLENSWSALSQIOIQTASTNNGQFETFPVLIN 240

Qy 241 AQONRVITITNDAGVVTNSIALLLNNNNMAAMDDVPMTQSGCGSYAL 289  
 Db 241 AQONRVITITNDAGVVTNSIALLLNNNNMAAMDDVPMTQSGCGSYAI 289

RESULT 5  
 AAR29272  
 ID AAR29272 standard; protein; 289 AA.  
 XX  
 AC AAR29272;  
 XX  
 FT Key Location/Qualifiers  
 FT Peptide 1..23  
 FT /label= sig\_peptide  
 FT /note= "putative N-terminal extension of the mature TCS"  
 FT Protein 24..270  
 FT /label= mat\_protein  
 FT Misc-difference 234  
 FT /note= "plant-derived TCS has Ser at this position"  
 FT Misc-difference 247  
 FT /note= "plant-derived TCS has Thr at this position"  
 FT Protein 270..289  
 FT /note= "putative C-terminal extension of the mature TCS"  
 FT  
 XX US5166056-A.  
 XX  
 PD 24-NOV-1992.  
 XX  
 XX 09-DEC-1991; 91US-00804293.  
 XX  
 XX 04-APR-1989; 89US-00333184.  
 XX  
 XX 07-SEP-1989; 89US-00404326.  
 XX  
 XX (GENE-) GENELABS INC.  
 XX  
 XX Piatak M, Chow TP;  
 XX  
 XX WPI; 1992-414954/50.  
 XX

PT Recombinant Trichosanthin protein prodn. in E. coli - for use in the  
 PT selective inhibition of viral expression in HIV infected cells.  
 XX  
 XX Disclosure; Fig 4; 37pp; English.  
 PS

The sequence is identical to that of plant-derived TCS except for two  
 CC conservative changes: a Thr for a Ser substitution at position 211 and a  
 CC Met for a Thr substitution at position 224. TCS is likely produced as a  
 CC secreted protein that undergoes post-translational processing at both the  
 CC amino and carboxy ends. The TCS coding sequence was amplified using the  
 CC primers of AAQ31828-30. The amplified prod. has the sequence of AAQ31827,  
 CC which was used as a probe. One isolate, pQ21D, comprises the sequence of  
 CC AAQ31826. The recombinant TCS sequence may be used in the recombinant  
 CC prodn. of TCS. TCS can be used for the selective inhibition of viral  
 CC expression in HIV-infected human T-cells or macrophages. (Updated on 25-  
 CC MAR-2003 to correct PF field.)  
 XX  
 SQ Sequence 289 AA;

Query Match 98.7%; Score 1413; DB 2; Length 289;  
 Best Local Similarity 98.6%; Pred. No. 1.9e-126;  
 Matches 285; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MIRFLVLSLLILFLTPPAVEGDVSRFGATSSSYGVFISNLRKALPNERKLYDIPLL 60  
 Db 1 MIRFLVLSLLILFLTPPAVEGDVSRFGATSSSYGVFISNLRKALPNERKLYDIPLL 60

Qy 61 RSSLPQSGQRYALIHNTYADETISVAIDVTNVYIMGYRAGDTSYFFNGASATEAAKYVFK 120  
 Db 61 RSSLPQSGQRYALIHNTYADETISVAIDVTNVYIMGYRAGDTSYFFNEASATEAAKYVFK 120

Qy 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPLGLPALDSAITTLFYNNANSASALMWLIQS 180  
 Db 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPLGLPALDSAITTLFYNNANSASALMWLIQS 180

Qy 181 TSEAAARYKFIEQQIGKRVDTFLPSLAIISLENSWSALSQIOIQTASTNNGQFESPVLIN 240  
 Db 181 TSEAAARYKFIEQQIGKRVDTFLPSLAIISLENSWSALSQIOIQTASTNNGQFETFPVLIN 240

Qy 241 AQONRVITITNDAGVVTNSIALLLNNNNMAAMDDVPMTQSGCGSYAL 289  
 Db 241 AQONRVITITNDAGVVTNSIALLLNNNNMAAMDDVPMTQSGCGSYAI 289

RESULT 6  
 AAR55129  
 ID AAR55129 standard; protein; 289 AA.  
 XX  
 AC AAR55129;  
 XX  
 DT 27-AUG-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 11-JAN-1995 (first entry)  
 XX  
 DE Alpha-trichosanthin coding.  
 XX  
 XX Virus; recombination; plant virus; alpha trichosanthin; phenotype;  
 KW alpha amylase; alpha haemoglobin; brome mosaic virus; gemini virus;  
 KW rice necrosis virus tobamovirus; gene expression; chinese cucumber.  
 XX  
 XX Trichosanthes kirilowii.  
 XX  
 XX US5316931-A.  
 XX  
 XX 31-MAY-1994.  
 XX  
 XX 31-JUL-1992; 92US-00923692.  
 XX  
 XX 26-FEB-1988; 88US-00160766.  
 XX  
 XX 26-FEB-1988; 88US-00160771.  
 XX  
 XX 15-JUL-1988; 88US-00219279.  
 XX  
 XX 17-FEB-1989; 89US-00310881.  
 XX  
 XX 05-MAY-1989; 89US-00347637.  
 XX



PR 08-JUN-1989; 89US-00363138.  
 PR 22-OCT-1980; 90US-00600244.  
 PR 16-JAN-1991; 91US-00641617.  
 PR 26-JUL-1991; 91US-00737899.  
 PR 01-AUG-1991; 91US-00739143.  
 XX  
 PA (BIOS-) BIOSOURCE GENETICS CORP.  
 XX  
 PI Garger SJ, Turpen AM, Grill LK, Grantham GL, Dawson WO, Donson J;  
 PI Turpen TH;  
 XX  
 DR WPI; 1994-176269/21.  
 DR N-PSDB; AAQ65573.  
 XX  
 PT New recombinant plant viral nucleic acid - capable of systemic infection  
 PT and stable expression of non-native nucleic acid in plant host.  
 XX  
 PS Example 4; Col 47-50; 44pp; English.  
 XX  
 CC The alpha-trichosanthin gene may be inserted into a recombinant plant  
 CC virus which can then be used to infect plants for the production of non-  
 CC native products (in this case alpha-trichosanthin). Other genes which  
 CC may be inserted into the virus are those which control a phenotypic  
 CC trait, such as male sterility, or sequences encoding anti-sense RNA which  
 CC can be useful to prevent the expression of undesired phenotypic traits.  
 CC The recombinant virus is derived from a plus sense, single stranded virus  
 CC selected from tobamovirus, bromo mosaic virus, rice necrosis virus or a  
 CC gemini virus. (Updated on 25-MAR-2003 to correct PF field.) (Updated on  
 CC 27-AUG-2003 to correct OS field.)  
 XX  
 SQ Sequence 289 AA;  
 Query Match 98.7%; Score 1413; DB 2; Length 289;  
 Best Local Similarity 98.6%; Pred. No. 1.9e-126;  
 Matches 285; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 MIRFLVLSLLILTLFTTPAVEGDVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60  
 Db 1 MIRFLVLSLLILTLFTTPAVEGDVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60  
 QY 61 RSSLPGSQRYALIHLTNYADETISVAIDVTNVIYMGYRAGDTSYFFNGASATEAAKYVFK 120  
 Db 61 RSSLPGSQRYALIHLTNYADETISVAIDVTNVIYMGYRAGDTSYFFNGASATEAAKYVFK 120  
 QY 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMVLIOQ 180  
 Db 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMVLIOQ 180  
 QY 181 TSEAAKYFIEQQIGKRVKDTFLPSLAIISLNSWSALSQIQIASTNNGQFESPVLIN 240  
 Db 181 TSEAAKYFIEQQIGKRVKDTFLPSLAIISLNSWSALSQIQIASTNNGQFESPVLIN 240  
 QY 241 AQORVTITNDAGVTSNIALLLNRNNMAAMDDVPMTQSFQCGSYAL 289  
 Db 241 AQORVTITNDAGVTSNIALLLNRNNMAAMDDVPMTQSFQCGSYAI 289  
 RESULT 7  
 AAW10468  
 ID AAW10468 standard; protein; 289 AA.  
 XX  
 AC AAW10468;  
 XX  
 DT 17-OCT-2003 (revised)  
 DT 26-APR-1997 (first entry)  
 XX  
 DE Chinese cucumber alpha-trichosanthin.  
 XX  
 KW Recombinant viral nucleic acid; RNA virus; vector; tobacco mosaic virus;  
 KW TMV; Chinese cucumber; alpha-trichosanthin;  
 KW ribosome inactivating protein; antiviral; virucide; transgenic plant.  
 XX  
 OS Trichosanthin kirilowii; Maximowicz.

XX WO9640867-A1.  
 XX  
 PD 19-DEC-1996.  
 XX  
 PF 06-JUN-1996; 96WO-US0009299.  
 XX  
 PR 07-JUN-1995; 95US-00483502.  
 XX  
 PA (BIOS-) BIOSOURCE TECHNOLOGIES INC.  
 XX  
 PI Donson J, Dawson WO, Grantham GL, Turpen TH, Turpen AM;  
 PI Garger SJ, Grill LK;  
 XX  
 DR WPI; 1997-065181/06.  
 DR N-PSDB; AAT47094.  
 XX  
 PT Recombinant viral nucleic acid producing, e.g. male sterility in plants -  
 PT comprises nucleic acid whose transcription is controlled by another  
 PT sequence.  
 XX  
 PS Example 4; Page 124-125; 149pp; English.  
 XX  
 CC Chinese cucumber alpha-trichosanthin (AAW10468) is a ribosome  
 CC inactivating protein of potential use in the treatment of HIV infection.  
 CC Expression vector pBG152 was constructed in which the alpha-  
 CC trichosanthin coding sequence (see also AAT47094) was placed under  
 CC control of the promoter of the tobacco mosaic virus-U1 coat protein gene,  
 CC which had been deleted. The viral nucleic acid was capable of self-  
 CC replication, encapsidation and systemic spread in infected Nicotiana  
 CC benthamiana plants, and directed the high-level expression of  
 CC biologically active alpha-trichosanthin in plant tissues. (Updated on 17-  
 CC OCT-2003 to standardise OS field)  
 XX  
 SQ Sequence 289 AA;  
 Query Match 98.7%; Score 1413; DB 2; Length 289;  
 Best Local Similarity 98.6%; Pred. No. 1.9e-126;  
 Matches 285; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 MIRFLVLSLLILTLFTTPAVEGDVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60  
 Db 1 MIRFLVLSLLILTLFTTPAVEGDVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60  
 QY 61 RSSLPGSQRYALIHLTNYADETISVAIDVTNVIYMGYRAGDTSYFFNGASATEAAKYVFK 120  
 Db 61 RSSLPGSQRYALIHLTNYADETISVAIDVTNVIYMGYRAGDTSYFFNGASATEAAKYVFK 120  
 QY 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMVLIOQ 180  
 Db 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMVLIOQ 180  
 QY 181 TSEAAKYFIEQQIGKRVKDTFLPSLAIISLNSWSALSQIQIASTNNGQFESPVLIN 240  
 Db 181 TSEAAKYFIEQQIGKRVKDTFLPSLAIISLNSWSALSQIQIASTNNGQFESPVLIN 240  
 QY 241 AQORVTITNDAGVTSNIALLLNRNNMAAMDDVPMTQSFQCGSYAL 289  
 Db 241 AQORVTITNDAGVTSNIALLLNRNNMAAMDDVPMTQSFQCGSYAI 289  
 RESULT 8  
 AAW11870  
 ID AAW11870 standard; protein; 289 AA.  
 XX  
 AC AAW11870;  
 XX  
 DT 27-AUG-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 21-APR-1997 (first entry)  
 XX  
 DE Chinese cucumber alpha-trichosanthin.  
 XX

KW	Recombinant virus; alpha-haemoglobin; human; chinese cucumber;	181	TSEARYKEIEQIGKRVKDTFLPLSLIISLNSWSALSQIQIASTNNQFETPVVLIN	240
KW	alpha-trichosanthin; rice; alpha amylase; beta-haemoglobin;	241	AQNQRVTITNDAGVVTSTNIALLLNNNMAAMDDVPMTQSCGGSYAL	289
KW	subgenomic promoter; coat protein.	241		289
OS	Cucumis sp.	241	AQNQRVTITNDAGVVTSTNIALLLNNNMAAMDDVPMTQSCGGSYAI	289
PN	US5589367-A.	RESULT 9		
PN	31-DEC-1996.	AAV01374		
XX	19-JAN-1994;	ID	AAV01374 standard; protein; 289 AA.	
XX	26-FEB-1988;	XX	AAV01374;	
XX	26-FEB-1988;	XX	20-MAR-2003 (revised)	
XX	15-JUL-1988;	DT	04-JUN-1999 (first entry)	
XX	17-FEB-1989;	XX	Chinese cucumber alpha-trichosanthin.	
XX	05-MAY-1989;	DE	Recombinant; plant virus; coat protein; systemic infection;	
XX	08-JUN-1989;	KW	transcription; therapeutic; chinese cucumber; alpha-trichosanthin.	
XX	22-OCT-1990;	XX	Cucumis sp.	
XX	16-JAN-1991;	XX	US5889190-A.	
XX	26-JUL-1991;	XX	30-MAR-1999.	
XX	01-AUG-1991;	XX	07-JUN-1995;	
XX	31-JUL-1992;	XX	26-FEB-1988;	
XX	(BIOS-) BIOSOURCE TECHNOLOGIES INC.	XX	26-FEB-1988;	
XX	Garger SJ, Turpen AM, Grill LK, Grantham GL, Donson J, Dawson WO;	XX	15-JUL-1988;	
PI	Turpen TH;	XX	17-FEB-1989;	
XX	WPI; 1997-076845/07.	XX	05-MAY-1989;	
DR	N-PSDB; AAT61376.	XX	08-JUN-1989;	
XX	Recombinant viral DNA for altering plant phenotype or protein prodn -	XX	22-OCT-1990;	
XX	contains non-native sub-genomic promoter for expression of heterologous	XX	16-JAN-1991;	
PT	protein and native promoter for expression of coat protein.	XX	26-JUL-1991;	
PT	Example 4; Col 45-46; 42pp; English.	XX	01-AUG-1991;	
PS	The sequences given in AAW11868-71 represent proteins which were produced	XX	31-JUL-1992;	
CC	by the recombinant viruses of the invention. The viruses are recombinant	XX	(BIOS-) BIOSOURCE TECHNOLOGIES INC.	
CC	plant viruses which comprise a native plant virus subgenomic promoter, at	XX	Garger SJ, Turpen AM, Grill LK, Grantham GL, Donson J, Dawson WO;	
CC	least one non-native plant virus subgenomic promoter, and a sequence	PI	Turpen TH;	
CC	encoding a plant virus coat protein. These heterologous sequences are	XX	WPI; 1997-076845/07.	
CC	preferably under the control of the native promoter sequence. By using a	DR	N-PSDB; AAT61376.	
CC	plant virus existing cells can be altered with a new coding sequences	XX	Recombinant viral DNA for altering plant phenotype or protein prodn -	
CC	without involving germ cell. The recombinant viruses are stable and can	XX	contains non-native sub-genomic promoter for expression of heterologous	
CC	cause systemic infection, with stable expression/transcription in plants	PT	protein and native promoter for expression of coat protein.	
CC	that are hosts for the non-native part of the vector. The nucleotide	PT	Example 4; Col 45-46; 42pp; English.	
CC	sequences encoding these protein preferably integrated in plant viruses	XX	The sequences given in AAW11868-71 represent proteins which were produced	
CC	having either the O-coat protein or the UI-coat protein gene. (Updated on	XX	by the recombinant viruses of the invention. The viruses are recombinant	
CC	25-MAR-2003 to correct PF field.) (Updated on 25-MAR-2003 to correct PR	XX	plant viruses which comprise a native plant virus subgenomic promoter, at	
CC	field.) (Updated on 27-AUG-2003 to correct OS field.)	XX	least one non-native plant virus subgenomic promoter, and a sequence	
XX	Sequence 289 AA;	XX	encoding a plant virus coat protein. These heterologous sequences are	
SQ	Query Match	XX	preferably under the control of the native promoter sequence. By using a	
	Best Local Similarity 98.7%; Score 1413; DB 2; Length 289;	XX	plant virus existing cells can be altered with a new coding sequences	
	Matches 285; Conservative 2; Mismatches 2; Indels 0; Gaps 0;	XX	without involving germ cell. The recombinant viruses are stable and can	
QY	1 MIRFLVLSLLILTLTTPAVEGDVSRFSGATSSSYGVFISNLRKALPNERKLYDIPLL	XX	cause systemic infection, with stable expression/transcription in plants	
Db	1 MIRFLVLSLLILTLTTPAVEGDVSRFSGATSSSYGVFISNLRKALPNERKLYDIPLL	XX	that are hosts for the non-native part of the vector. The nucleotide	
QY	61 RSLPQSQRYALHLTNVADETTSVADITNVIMYGRAGTSSVFENGASATEAAKVFK	XX	sequences encoding these protein preferably integrated in plant viruses	
Db	61 RSLPQSQRYALHLTNVADETTSVADITNVIMYGRAGTSSVFENGASATEAAKVFK	XX	having either the O-coat protein or the UI-coat protein gene. (Updated on	
QY	121 DAMRKVTLPSYGNVRLQTAGKIRENIPGLPALDSAITTLFYNNANSAASALMWLIQS	XX	25-MAR-2003 to correct PF field.) (Updated on 25-MAR-2003 to correct PR	
Db	121 DAMRKVTLPSYGNVRLQTAGKIRENIPGLPALDSAITTLFYNNANSAASALMWLIQS	XX	field.) (Updated on 27-AUG-2003 to correct OS field.)	
QY	181 TSEARYKEIEQIGKRVKDTFLPLSLIISLNSWSALSQIQIASTNNQFETPVVLIN	XX	Sequence 289 AA;	
QY		XX	Query Match	
		XX	Best Local Similarity 98.7%; Score 1413; DB 2; Length 289;	
		XX	Matches 285; Conservative 2; Mismatches 2; Indels 0; Gaps 0;	

XX SQ Sequence 289 AA;

Query Match 98.7%; Score 1413; DB 2; Length 289;  
 Best Local Similarity 98.6%; Pred. No. 1.9e-126;  
 Matches 285; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MIRFLVLSLLILTLFTTPAVEGDVSRFSLGATSSSYGVFISNLRKALPNERKLYDIPLL 60  
 Db 1 MIRFLVLSLLILTLFTTPAVEGDVSRFSLGATSSSYGVFISNLRKALPNERKLYDIPLL 60

QY 61 RSSLPGSQRYALIHLYNADETISVAIDVTNNVIMGYRAGDTSYFNGASATEAAKYVFK 120  
 Db 61 RSSLPGSQRYALIHLYNADETISVAIDVTNNVIMGYRAGDTSYFNGASATEAAKYVFK 120

QY 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAALMWLIQS 180  
 Db 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAALMWLIQS 180

QY 181 TSEAAKYKFIQQIGKRVDTFLPSLAIIISLNSWSALSQIQIASTNNQFETPVVLIN 240  
 Db 181 TSEAAKYKFIQQIGKRVDTFLPSLAIIISLNSWSALSQIQIASTNNQFETPVVLIN 240

QY 241 AQNRVMTITNDAGVVTSNIAILLNRRNMAAMDDDDVPMTQSGCGSYAL 289  
 Db 241 AQNRVMTITNDAGVVTSNIAILLNRRNMAAMDDDDVPMTQSGCGSYAI 289

RESULT 10  
 AAW84192  
 ID AAW84192 standard; protein; 289 AA.

XX AC AAW84192;  
 XX 27-AUG-2003 (revised)  
 DT 01-APR-1999 (first entry)  
 XX Chinese cucumber alpha-trichosanthin.

XX Chinese cucumber; alpha-trichosanthin; plant virus;  
 KW RNA plant virus promoter; systemic infection; foreign gene expression;  
 KW AIDS therapeutic drug.

XX Momordica cochinchinensis.  
 OS US5866785-A.  
 EN 02-FEB-1999.  
 XX 07-JUN-1995; 95US-00482920.  
 XX 26-FEB-1988; 88US-00160766.  
 PR 26-FEB-1988; 88US-00160771.  
 PR 15-JUL-1988; 88US-00219279.  
 PR 17-FEB-1989; 89US-00310881.  
 PR 05-MAY-1989; 89US-00347637.  
 PR 08-JUN-1989; 89US-00363138.  
 PR 22-OCT-1990; 90US-00600244.  
 PR 16-JAN-1991; 91US-00641617.  
 PR 26-JUL-1991; 91US-00737899.  
 PR 01-AUG-1991; 91US-00739143.  
 PR 31-JUL-1992; 92US-00923692.  
 PR 19-JAN-1994; 94US-00184237.  
 XX (BIOS-) BIOSOURCE TECHNOLOGIES INC.

XX Garger SJ, Turpen AM, Grill LK, Grantham GL, Donson J, Dawson WO;  
 PI Turpen TH;  
 XX WPI; 1999-142035/12.  
 DR N-PSDB; AAX03385.  
 XX Recombinant plant viral vector - that is capable of systemic infection in

PT host plant and stable production of heterologous DNA useful for producing  
 XX therapeutic proteins for treating e.g. AIDS.  
 PS Example 4; Col 45-48; 45pp; English.  
 XX The present sequence represents chinese cucumber alpha-trichosanthin. The  
 CC nucleic acid sequence can be expressed in the plant viral constructs of  
 CC the invention. The specification describes a recombinant plant viral  
 CC nucleic acid derived from a positive (+)-sense RNA plant virus comprising  
 CC a native (+)-sense RNA plant virus promoter that is linked to an  
 CC expression sequence and a heterologous (+)-sense RNA plant virus promoter  
 CC that is linked to an expression sequence. The promoters are incapable of  
 CC recombination with each other, and one of the expression sequences  
 CC encodes a plant viral coat protein while the other is optionally a  
 CC heterologous coding sequence. The plant viral nucleic acid is capable of  
 CC systemic infection in a host plant. The viral construct is useful for the  
 CC introduction and expression of non-viral foreign genes in plants and the  
 CC production of e.g. potential AIDS therapeutic drugs. (Updated on 27-AUG-  
 CC 2003 to correct Os field.)  
 XX Sequence 289 AA;  
 SQ

Query Match 98.7%; Score 1413; DB 2; Length 289;  
 Best Local Similarity 98.6%; Pred. No. 1.9e-126;  
 Matches 285; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MIRFLVLSLLILTLFTTPAVEGDVSRFSLGATSSSYGVFISNLRKALPNERKLYDIPLL 60  
 Db 1 MIRFLVLSLLILTLFTTPAVEGDVSRFSLGATSSSYGVFISNLRKALPNERKLYDIPLL 60

QY 61 RSSLPGSQRYALIHLYNADETISVAIDVTNNVIMGYRAGDTSYFNGASATEAAKYVFK 120  
 Db 61 RSSLPGSQRYALIHLYNADETISVAIDVTNNVIMGYRAGDTSYFNGASATEAAKYVFK 120

QY 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAALMWLIQS 180  
 Db 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAALMWLIQS 180

QY 181 TSEAAKYKFIQQIGKRVDTFLPSLAIIISLNSWSALSQIQIASTNNQFETPVVLIN 240  
 Db 181 TSEAAKYKFIQQIGKRVDTFLPSLAIIISLNSWSALSQIQIASTNNQFETPVVLIN 240

QY 241 AQNRVMTITNDAGVVTSNIAILLNRRNMAAMDDDDVPMTQSGCGSYAL 289  
 Db 241 AQNRVMTITNDAGVVTSNIAILLNRRNMAAMDDDDVPMTQSGCGSYAI 289

RESULT 11  
 AAY87791  
 ID AAY87791 standard; protein; 289 AA.  
 XX AC AAY87791;  
 XX 06-AUG-2003 (revised)  
 DT 24-AUG-2000 (first entry)  
 XX Chinese cucumber alpha-trichosanthin protein.

XX Animal RNA virus; viral coat protein; plant; male sterility; interleukin;  
 KW EPO; erythropoietin; CSF; colony stimulating factor; Factor VIII; hGH;  
 KW human growth hormone; melanin; insulin; vaccine;  
 KW stereo specific catalysis; alpha-trichosanthin.  
 XX Trichosanthin kirilowii.  
 OS US6054566-A.  
 EN 25-APR-2000.  
 XX 07-JUN-1995; 95US-00484341.  
 PR 26-FEB-1988; 88US-00160766.  
 PR 26-FEB-1988; 88US-00160771.

27-AUG-2003 (revised)  
25-MAR-2003 (revised)  
17-JUN-1993 (first entry)  
Encodes chinese cucumber alpha-tricosanthin.  
Recombinant products; commercial production; fermentation; biosynthesis;  
natural products; recombinant proteins; product expression;  
protein expression; expressed proteins.  
Cucurbitaceae.  
WO9303161-A1.  
18-FEB-1993.  
31-JUL-1992; 92WO-US006359.  
01-AUG-1991; 91US-00739143.  
(DONS// DONSON J.  
(DAWS// DAWSON W O.  
(GRAN// GRANTHAM G L.  
(TURP// TURPEN T H.  
(TURP// TURPEN A M.  
(GARG// GARGER S J.  
(GRIL// GRILLE L K.  
Donson J, Dawson WO, Grantham GL, Turpen TH, Turpen AM;  
Garger SJ, Grille LK;  
WPI; 1993-076518/09.  
N-PSDB; AAQ37679.  
Recombinant plant viral nucleic acids - used to express a prod., e.g.  
antibody or IL-1 in a plant.  
Example 4; Page 96; 30pp; English.  
This sequence represents chinese cucumber alpha-tricosanthin. The coding  
sequence is inserted into a recombinant plant viral nucleic acid which is  
then used to express a recombinant product (in this case alpha-  
tricosanthin) in a plant. The plant viral sequence may be from tobacco  
mosaic, cucumber green mottle, cowpea mosaic, bromo mosaic, broad bean  
mottle, rice necrosis, geminiviruses, tomato golden mosaic, Cassava  
latent and maize streak viruses. (Updated on 25-MAR-2003 to correct PN  
field.) (Updated on 25-MAR-2003 to correct PA field.) (Updated on 25-MAR-  
2003 to correct PI field.) (Updated on 27-AUG-2003 to correct OS field.)  
Sequence 289 AA;  
Query Match 98.5%; Score 1410; DB 2; Length 289;  
Best Local Similarity 98.3%; Pred. No. 3.6e-126;  
Matches 284; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
1 MIRFLVLSLLILTLFTTPAVEGDVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60  
1 MIRFLVLSLLILTLFTTPAVEGDVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60  
61 RSSLPGSQRYALIHLYTNYADETISVAIDVTNVIYMGYRAGDTSYFFNGASATEAAKYVFK 120  
61 RSSLPGSQRYALIHLYTNYADETISVAIDVTNVIYMGYRAGDTSYFFNGASATEAAKYVFK 120  
121 DAMRKVTLPSGNYERLQTAAGKIRENIPLGLPALDSAITTLFYNNANSAASALMWLIQS 180  
121 DAMRKVTLPSGNYERLQTAAGKIRENIPLGLPALDSAITTLFYNNANSAASALMWLIQS 180  
181 TSEAAKYKFEQIQIGKRVKDTFLPSLAISLENSWSALSQKIQIASTNNQGFETPVVLIN 240  
181 TSEAAKYKFEQIQIGKRVKDTFLPSLAISLENSWSALSQKIQIASTNNQGFETPVVLIN 240  
241 AQQRVITITNDAGVVTSTNIALLLNRNNMAAMDDVPMTQSGCGSYAL 289  
AQQRVITITNDAGVVTSTNIALLLNRNNMAAMDDVPMTQSGCGSYAL 289

15-JUL-1988; 88US-00219279.  
17-FEB-1989; 89US-00310881.  
05-MAY-1989; 89US-00347637.  
08-JUN-1989; 89US-00363138.  
22-OCT-1990; 90US-00600244.  
16-JAN-1991; 91US-00641617.  
26-JUN-1991; 91US-00737899.  
01-AUG-1991; 91US-00739143.  
31-JUL-1992; 92US-00923692.  
(BIOS-) BIOSOURCE TECHNOLOGIES INC.  
Garger SJ, Grill LK, Turpen TH, Grantham GL, Dawson WO;  
Turpen AM, Donson J;  
N-PSDB; AAA12393.  
Recombinant non-retroviral nucleic acid for producing proteins such as  
interleukins, melanin and vaccines, comprises subgenomic promoters linked  
to sequences coding for viral coat protein and heterologous proteins.  
Example 4; Col 61-62; 51pp; English.  
This invention describes a novel recombinant viral nucleic acid (I) from  
a non-retroviral (+) sense, single stranded animal RNA virus comprising a  
nucleic acid sequence coding for a viral coat protein regulated by a  
native subgenomic promoter and other two heterologous nucleic acid  
sequences regulated by two other subgenomic promoters. (I) is useful for  
expressing foreign genes e.g. genes inducing male sterility in plants.  
(I) is also useful for producing proteins such as interleukins, EPO  
(erythropoietin), CSF (colony stimulating factor), Factor VIII, hGH  
(human growth hormone), melatonin, insulin, vaccines etc., and enzymes that  
are useful for stereo specific catalysis of organic compounds. (I) is  
stable and transcribed systemically. The dual subgenomic promoter system  
reduces the frequency of recombination thus reducing regeneration of the  
wild type virus. This sequence represents a chinese cucumber alpha-  
tricosanthin protein which is described in the method of the invention.  
(Updated on 06-AUG-2003 to correct OS field.)  
Sequence 289 AA;  
Query Match 98.7%; Score 1413; DB 3; Length 289;  
Best Local Similarity 98.6%; Pred. No. 1.9e-126;  
Matches 285; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
1 MIRFLVLSLLILTLFTTPAVEGDVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60  
1 MIRFLVLSLLILTLFTTPAVEGDVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60  
61 RSSLPGSQRYALIHLYTNYADETISVAIDVTNVIYMGYRAGDTSYFFNGASATEAAKYVFK 120  
61 RSSLPGSQRYALIHLYTNYADETISVAIDVTNVIYMGYRAGDTSYFFNGASATEAAKYVFK 120  
121 DAMRKVTLPSGNYERLQTAAGKIRENIPLGLPALDSAITTLFYNNANSAASALMWLIQS 180  
121 DAMRKVTLPSGNYERLQTAAGKIRENIPLGLPALDSAITTLFYNNANSAASALMWLIQS 180  
181 TSEAAKYKFEQIQIGKRVKDTFLPSLAISLENSWSALSQKIQIASTNNQGFETPVVLIN 240  
181 TSEAAKYKFEQIQIGKRVKDTFLPSLAISLENSWSALSQKIQIASTNNQGFETPVVLIN 240  
241 AQQRVITITNDAGVVTSTNIALLLNRNNMAAMDDVPMTQSGCGSYAL 289  
241 AQQRVITITNDAGVVTSTNIALLLNRNNMAAMDDVPMTQSGCGSYAL 289  
RESULT 12  
AAR32986  
ID AAR32986 standard; protein; 289 AA.  
XX AAR32986;  
AC AAR32986;  
XX AAR32986;

Db 241 AQRVMTITNDAGVTSNTIALLLNRNNMAAMDDDDVPMWTSFGCGTYAI 289

## RESULT 13

AAR07523

ID AAR07523 standard; protein; 289 AA.

XX AC AAR07523;

DT 06-FEB-1991 (first entry)

XX Alpha-Trichosanthin encoded by insert sequence from clone 12.

XX trichosanthin (TCS); Human Immunodeficiency Virus; inhibition.

XX Trichosanthin Kirilowii.

XX Key Location/Qualifiers

FT Region 1..23

FT Protein /label= signal peptide

FT FT 24..289

FT FT /label= alpha-TCS

XX W09012097-A.

XX 18-OCT-1990.

XX 04-APR-1989; 89US-00333184.

XX 04-APR-1989; 89US-00333184.

XX (GENE-) GENELABS INC.

XX Piatek M, Chow T, Fry K;

XX WPI; 1990-334847/44.

XX N-PSDB; AAQ06351.

XX Recombinant trichosanthin protein - with selective inhibitory effect on viral expression in HIV infected T-cells or monocyte-macrophage.

XX Example; Fig 20; 102pp; English.

XX PQ30E was used as a probe to identify clones containing sequences coding for TCS in a T.Kirilowii genomic library. Five clones were eventually isolated and sequenced, including clone 12. They were found to have homology to the alpha-TCS coding sequence. See also AAQ06343-Q06350

XX Sequence 289 AA;

Query Match 92.6%; Score 1326; DB 2; Length 289;  
Best Local Similarity 92.4%; Pred. No. 3.9e-118;  
Matches 267; Conservative 8; Mismatches 14; Indels 0; Gaps 0;

Qy 1 MIRELVSLILTLTTPAVEGDSFRLSGATSSSYGVFISNRKALPNERKYDIPLL 60

Db 1 MIRELVSLILTLTTPAVEGDSFRLSGATSSSYGVFISNRKALPYERRLYDISLL 60

Qy 61 RSSLPQSQRYALIHLYTNYADETISVAIDVTNVIYMGVAGDTSYFFNGASATEAAKYVFK 120

Db 61 RSTLQSQRYALIHLYTNYADETISVAIDVTNVIYMGVAGDTSYFFNEASATEAAKYVFK 120

Qy 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAALMWLIQS 180

Db 121 DAQRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAALMWLIQS 180

Qy 181 TSEAAKYKFIHQOIGKRVKDTFLPSLAISLNSWSALSQIQIASTNNQGFSPVVLIN 240

Db 181 MSEAARYKFIHQOIGKRVKDTFLPSLAISLNSWSALSQIQIASTNNQGFSPVVLIN 240

Qy 241 AQRVMTITNDAGVTSNTIALLLNRNNMAAMDDDDVPMWTSFGCGTYAI 289

Db 241 AQRVMTITNDAGVTSNTIALLLNRNNMAVIDDHPVMAQSFSGGSYAI 289

## RESULT 14

AAR29276

ID AAR29276 standard; protein; 289 AA.

XX AC AAR29276;

DT 25-MAR-2003 (revised)

DT 13-JAN-1993 (first entry)

XX Ribosome inactivating protein deduced from pQ12.

XX TCS; alpha-trichosanthin; Radix trichosanthin; abortifacient; RIP;

XX multi-gene family; Maxim; HIV-infected human T cells;

XX human immunodeficiency virus; probe.

XX Trichosanthin Kirilowii.

XX Key Location/Qualifiers

FT Peptide 1..23

FT Protein /label= signal

XX 24..289

XX US5128460-A.

XX 07-JUL-1992.

XX 04-APR-1990; 90US-00504775.

XX 04-APR-1989; 89US-00333184.

XX 07-SEP-1989; 89US-00404326.

XX (GENE-) GENELABS INC.

XX Piatek M, Chow TP, Fry K;

XX WPI; 1992-249485/30.

XX N-PSDB; AAQ26505.

XX Nucleic acid encoding trichosanthin protein - which can be used to inactivate ribosome(s) to inhibit protein synthesis or to inhibit HIV expression.

XX Example 7; Fig 20; 53pp; English.

XX The clone PQ30E, containing a 0.6kb insert, which was isolated from a T.Kirilowii library (see AAQ26501), was used as a probe to isolate additional sequences homologous to alpha-TCS, i.e. other members of the RIP multi-gene family. Four unique cloned inserts were identified and designated pQ2, pQ3, pQ12 and pQ24. Each insert encoded a putative full length RIP protein. See AAQ26499-Q26508. (Updated on 25-MAR-2003 to correct PF field.) (Updated on 25-MAR-2003 to correct PR field.)

XX Sequence 289 AA;

Query Match 92.6%; Score 1326; DB 2; Length 289;  
Best Local Similarity 92.4%; Pred. No. 3.9e-118;  
Matches 267; Conservative 8; Mismatches 14; Indels 0; Gaps 0;

Qy 1 MIRELVSLILTLTTPAVEGDSFRLSGATSSSYGVFISNRKALPNERKYDIPLL 60

Db 1 MIRELVSLILTLTTPAVEGDSFRLSGATSSSYGVFISNRKALPYERRLYDISLL 60

Qy 61 RSSLPQSQRYALIHLYTNYADETISVAIDVTNVIYMGVAGDTSYFFNGASATEAAKYVFK 120

Db 61 RSTLQSQRYALIHLYTNYADETISVAIDVTNVIYMGVAGDTSYFFNEASATEAAKYVFK 120

Qy 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAALMWLIQS 180

Db 121 DAQRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAALMWLIQS 180

Qy 181 TSEAAKYKFIHQOIGKRVKDTFLPSLAISLNSWSALSQIQIASTNNQGFSPVVLIN 240

QY 23 GDVSRFLSGATSSSYGVFISNLRKALPNERKLYDIPLRRSSLPQSORYALIHILTNVADET 82  
Db 1 GDVSRFLSGATSSSYGVFISNLRKALPNERKLYDIPLRRSSLPQSORYALIHILTNVADET 60  
QY 83 ISVAIDVTNVIYIMGYRAGDTSYFFNGASATEAAKYVFKDAMRKVTLPYSGNYERLQTAAG 142  
Db 61 ISVAIDVTNVIYIMGYRAGDTSYFFNEASATEAAKYVFKDAMRKVTLPYSGNYERLQTAAG 120  
QY 143 KIRENIPGLPALDLSAITTLFYNNANSAASALMVLIOSTSEAAARYKFEIOQIGKRVDTKF 202  
Db 121 KIRENIPGLPALDLSAITTLFYNNANSAASALMVLIOSTSEAAARYKFEIOQIGKRVDTKF 180  
QY 203 LPSLAIIISLENSWSALSQIQIASTNNGQFESPVWLINAQNRVTITNVDAGVVTSNIAL 262  
Db 181 LPSLAIIISLENSWSALSQIQIASTNNGQFESPVWLINAQNRVTITNVDAGVVTSNIAL 240  
QY 263 LLNRNNMAAMDDDDVPMTQSFQCGSYAL 289  
Db 241 LLNRNNMAAMDDDDVPMTQSFQCGSYAI 267

Search completed: September 10, 2004, 13:58:34  
Job time : 124 secs

Db 181 MSAAARYKTEIQIGRVRDKTFLPSLAIIISLENSWSALSQIQIASTNNGQFETPVVLIN 240  
QY 241 AQQRVTITNVDAGVVTSNIALLLNRNNMAAMDDDDVPMTQSFQCGSYAL 289  
Db 241 AQQRVTITNVDAGVVTSNIALLLNRNNMAVIDDHVPMAQSFQCGSYAI 289  
RESULT 15  
AAW21703  
ID AAW21703 standard; protein; 267 AA.  
XX  
AC AAW21703;  
XX 25-MAR-2003 (revised)  
DT 26-SEP-1997 (first entry)  
XX  
DE Trichosanthin.  
XX pro-Ribosome Inactivating Protein; proRIP; peptide linker; cancer;  
KW inactivation; eukaryotic ribosome; alpha fragment; beta fragment;  
KW inhibitor; protein synthesis; N-glycosidase; glycosidic bond; liver; rat;  
KW ribosomal 28S RNA; cellular proliferation; HIV-infected T cell.  
XX  
OS Trichosanthes kirilowii.

XX Key Location/Qualifiers  
FH 139. .149  
FT Region /note= "Position of possible insertion of internal  
FT peptide linker sequence"  
XX  
XX US5635384-A.  
XX  
XX 03-JUN-1997.  
XX  
XX 26-JAN-1995; 95US-00378761.  
XX  
XX 11-JUN-1990; 90US-00535636.  
XX 09-DEC-1992; 92US-00987927.  
XX  
XX (DOWC ) DOWELANCO.  
XX  
XX Hey TD, Morgan ABR, Walsh TA;  
XX  
XX WPI; 1997-309831/28.  
XX  
XX Inactive precursor of maize ribosome-inactivating protein - also chimeric  
XX ribosome-inactivating protein precursors containing internal linker  
XX sequences.  
XX  
XX Claim 2; Col 115-118; 121pp; English.

XX The sequences given in AAW21698-710 represent Ribosome Inactivating  
XX Proteins (RIP's), which may be used in the construction of the proRIP of  
XX the invention. The proRIP has a selectively removable, internal peptide  
XX linker. The precursor sequence is incapable of inactivating eukaryotic  
XX ribosomes, but can be converted by removal of the linker into a protein  
XX having alpha and beta fragments and being capable of inactivating  
XX eukaryotic ribosomes. RIPs are potent inhibitors of eukaryotic protein  
XX synthesis. They possess a highly specific N-glycosidase activity which  
XX cleaves the glycosidic bond of adenine 4324 of rat liver ribosomal 28S  
XX RNA. RIP's selectively inhibit cellular proliferation of cells, e.g.  
XX cancer cells and HIV-infected T cells. The inactive proRIP proteins make  
XX it possible to provide protein synthesis inhibitors with uses in  
XX practical and improved ways not before possible. The RIP can be used to  
XX make cytotoxic conjugates. (Updated on 25-MAR-2003 to correct PF field.)

XX SQ Sequence 267 AA;  
Query Match 92.2%; Score 1321; DB 2; Length 267;  
Best Local Similarity 99.3%; Pred. No. 1e-117;  
Matches 265; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 13:59:12 ; Search time 138 Seconds  
(without alignments)  
671.589 Million cell updates/sec

Title: US-09-905-247-1

Perfect score: 1432

Sequence: 1 MIRFLVLSLLTLFLTPA.....AAMDDVPMQTSGCGSYAL 289

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1335176 seqs, 320689617 residues

Total number of hits satisfying chosen parameters: 1335176

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA.\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1413	98.7	289	12	US-10-280-725B-4
2	1413	98.7	289	14	US-10-280-679B-4
3	1214	84.8	247	9	US-09-792-793A-39
4	1214	84.8	247	15	US-10-375-209A-39
5	1202	83.9	247	14	US-10-127-890-6
6	1064	74.3	247	9	US-09-792-793A-34
7	1064	74.3	247	15	US-10-375-209A-34
8	806.5	56.3	263	14	US-10-127-890-7
9	717.5	50.1	248	14	US-10-127-890-5
10	703	49.1	263	14	US-10-127-890-4
11	419	29.3	267	14	US-10-282-935-1
12	419	29.3	267	14	US-10-127-890-1
13	419	29.3	267	15	US-10-440-796-1
14	419	29.3	576	14	US-10-083-336A-1
15	340	23.7	198	14	US-10-083-336A-3

16	340	23.7	198	14	US-10-083-336A-7	Sequence 7, Appli
17	340	23.7	199	14	US-10-083-336A-5	Sequence 5, Appli
18	340	23.7	200	14	US-10-083-336A-10	Sequence 10, Appli
19	335.5	23.4	185	14	US-10-083-336A-9	Sequence 9, Appli
20	334	23.3	188	14	US-10-083-336A-4	Sequence 4, Appli
21	334	23.3	188	14	US-10-083-336A-8	Sequence 8, Appli
22	334	23.3	189	14	US-10-083-336A-6	Sequence 6, Appli
23	334	23.3	190	14	US-10-083-336A-11	Sequence 3, Appli
24	327	22.8	251	14	US-10-282-935-3	Sequence 11, Appli
25	327	22.8	251	15	US-10-440-796-3	Sequence 3, Appli
26	327	22.8	293	9	US-09-765-527-259	Sequence 259, App
27	327	22.8	309	9	US-09-765-527-253	Sequence 253, App
28	327	22.8	332	9	US-09-765-527-251	Sequence 251, App
29	323	22.6	252	9	US-09-347-064-2	Sequence 2, Appli
30	323	22.6	252	9	US-09-347-064-8	Sequence 8, Appli
31	320	22.3	251	14	US-10-127-890-99	Sequence 99, Appli
32	320	22.3	251	14	US-10-127-890-101	Sequence 101, App
33	320	22.3	251	14	US-10-127-890-107	Sequence 107, App
34	319	22.3	251	14	US-10-127-890-110	Sequence 110, App
35	319	22.3	251	14	US-10-127-890-111	Sequence 111, App
36	319	22.3	316	14	US-10-127-890-1	Sequence 1, Appli
37	318.5	22.2	507	14	US-10-074-596-1	Sequence 11, Appli
38	318	22.2	251	9	US-09-765-527-247	Sequence 247, App
39	318	22.2	251	14	US-10-127-890-2	Sequence 2, Appli
40	318	22.2	251	14	US-10-127-890-100	Sequence 100, App
41	318	22.2	251	14	US-10-127-890-102	Sequence 102, App
42	318	22.2	251	14	US-10-127-890-103	Sequence 103, App
43	318	22.2	251	14	US-10-127-890-105	Sequence 105, App
44	318	22.2	251	14	US-10-127-890-109	Sequence 109, App
45	317	22.1	251	14	US-10-127-890-106	Sequence 106, App

## ALIGNMENTS

### RESULT 1

US-10-280-725B-4  
; Sequence 4, Application US/10280725B  
; Publication No. US20040049025A1  
; GENERAL INFORMATION:  
; APPLICANT: Large Scale Biology Corporation  
; TITLE OF INVENTION: Recombinant Viral Nucleic Acids  
; FILE REFERENCE: LSEC-0109-US02  
; CURRENT APPLICATION NUMBER: US/10/280,725B  
; CURRENT FILING DATE: 2002-10-24  
; PRIOR APPLICATION NUMBER: 09/557,941  
; PRIOR FILING DATE: 2000-04-24  
; PRIOR APPLICATION NUMBER: 08/484,341  
; PRIOR FILING DATE: 1995-06-07  
; PRIOR APPLICATION NUMBER: 07/923,692  
; PRIOR FILING DATE: 1992-07-31  
; PRIOR APPLICATION NUMBER: 07/600,244  
; PRIOR FILING DATE: 1990-10-22  
; PRIOR APPLICATION NUMBER: 07/641,617  
; PRIOR FILING DATE: 1991-01-16  
; PRIOR APPLICATION NUMBER: 07/737,899  
; PRIOR FILING DATE: 1991-07-26  
; PRIOR APPLICATION NUMBER: 07/739,143  
; PRIOR FILING DATE: 1991-08-01  
; PRIOR APPLICATION NUMBER: 07/310,881  
; PRIOR FILING DATE: 1989-02-17  
; PRIOR APPLICATION NUMBER: 07/160,766  
; PRIOR FILING DATE: 1988-02-26  
; PRIOR APPLICATION NUMBER: 07/160,771  
; PRIOR FILING DATE: 1988-02-26  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 289  
; TYPE: PRT  
; ORGANISM: Chinese cucumber  
US-10-280-725B-4

Query Match 98.7%; Score 1413; DB 12; Length 289;  
 Best Local Similarity 98.6%; Pred. No. 3.7e-134;  
 Matches 285; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MIRFLVLSLLILTLFTTAVAGDVSRFLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60  
 DB 1 MIRFLVLSLLILTLFTTAVAGDVSRFLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60

QY 61 RSSLPQSQRVYALHILTNVADETISVAIDVTNVYIMGYRAGDTSYFFNGASATEAAKYVEK 120  
 DB 61 RSSLPQSQRVYALHILTNVADETISVAIDVTNVYIMGYRAGDTSYFFNEASATEAAKYVEK 120

QY 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPLGLPALDSAITTLFYNNANSAASALMVLIO 180  
 DB 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPLGLPALDSAITTLFYNNANSAASALMVLIO 180

QY 181 TSEAAKYKFEIQOIGKRVKDTFLPSLAIISLNSWSALSQIQIASTNNGQFETPVVLIN 240  
 DB 181 TSEAAKYKFEIQOIGKRVKDTFLPSLAIISLNSWSALSQIQIASTNNGQFETPVVLIN 240

QY 241 AQNRVTITNVDAAGVTSNTALLNRNMAAMDDVPMTQSGCGSYAL 289  
 DB 241 AQNRVTITNVDAAGVTSNTALLNRNMAAMDDVPMTQSGCGSYAI 289

RESULT 2

US-10-280-679B-4  
 ; Sequence 4, Application US/10280679B  
 ; Publication No. US20030150019A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Large Scale Biology Corporation  
 ; TITLE OF INVENTION: Monopartite RNA Virus Transformation Vectors  
 ; FILE REFERENCE: LSPC-0109-US03  
 ; CURRENT APPLICATION NUMBER: US/10/280,679B  
 ; CURRENT FILING DATE: 2000-10-24  
 ; PRIOR APPLICATION NUMBER: 09/557,941  
 ; PRIOR FILING DATE: 2000-04-24  
 ; PRIOR APPLICATION NUMBER: 08/484,341  
 ; PRIOR FILING DATE: 1995-06-07  
 ; PRIOR APPLICATION NUMBER: 07/923,692  
 ; PRIOR FILING DATE: 1992-07-31  
 ; PRIOR APPLICATION NUMBER: 07/600,244  
 ; PRIOR FILING DATE: 1990-10-22  
 ; PRIOR APPLICATION NUMBER: 07/641,617  
 ; PRIOR FILING DATE: 1991-01-16  
 ; PRIOR APPLICATION NUMBER: 07/737,899  
 ; PRIOR FILING DATE: 1991-07-26  
 ; PRIOR APPLICATION NUMBER: 07/739,143  
 ; PRIOR FILING DATE: 1991-08-01  
 ; PRIOR APPLICATION NUMBER: 07/310,881  
 ; PRIOR FILING DATE: 1989-02-17  
 ; PRIOR APPLICATION NUMBER: 07/160,766  
 ; PRIOR FILING DATE: 1988-02-26  
 ; PRIOR APPLICATION NUMBER: 07/160,771  
 ; PRIOR FILING DATE: 1988-02-26  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 11  
 ; SOFTWARE: Patentin version 3.1  
 ; SEQ ID NO 4  
 ; LENGTH: 289  
 ; TYPE: PRT  
 ; ORGANISM: Chinese cucumber protein alpha-trichosanthin  
 US-10-280-679B-4

Query Match 98.7%; Score 1413; DB 14; Length 289;  
 Best Local Similarity 98.6%; Pred. No. 3.7e-134;  
 Matches 285; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MIRFLVLSLLILTLFTTAVAGDVSRFLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60  
 DB 1 MIRFLVLSLLILTLFTTAVAGDVSRFLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60

QY 61 RSSLPQSQRVYALHILTNVADETISVAIDVTNVYIMGYRAGDTSYFFNGASATEAAKYVEK 120  
 DB 61 RSSLPQSQRVYALHILTNVADETISVAIDVTNVYIMGYRAGDTSYFFNEASATEAAKYVEK 120

QY 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPLGLPALDSAITTLFYNNANSAASALMVLIO 180  
 DB 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPLGLPALDSAITTLFYNNANSAASALMVLIO 180

QY 181 TSEAAKYKFEIQOIGKRVKDTFLPSLAIISLNSWSALSQIQIASTNNGQFETPVVLIN 240  
 DB 181 TSEAAKYKFEIQOIGKRVKDTFLPSLAIISLNSWSALSQIQIASTNNGQFETPVVLIN 240

QY 241 AQNRVTITNVDAAGVTSNTALLNRNMAAMDDVPMTQSGCGSYAL 289  
 DB 241 AQNRVTITNVDAAGVTSNTALLNRNMAAMDDVPMTQSGCGSYAI 289

RESULT 3

US-09-792-793A-39  
 ; Sequence 39, Application US/09792793A  
 ; Patent No. US20020168370A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: McDonald, John R.  
 ; APPLICANT: Coggin, Philip  
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING SECONDARY TISSUE DAMAGE AND  
 ; TITLE OF INVENTION: OTHER INFLAMMATORY CONDITIONS AND DISORDERS  
 ; FILE REFERENCE: 25020-601D  
 ; CURRENT APPLICATION NUMBER: US/09/792,793A  
 ; CURRENT FILING DATE: 2001-02-22  
 ; NUMBER OF SEQ ID NOS: 93  
 ; SOFTWARE: Patentin Ver. 2.0  
 ; SEQ ID NO 39  
 ; LENGTH: 247  
 ; TYPE: PRT  
 ; ORGANISM: Trichosanthens kirilowii  
 US-09-792-793A-39

Query Match 84.8%; Score 1214; DB 9; Length 247;  
 Best Local Similarity 99.6%; Pred. No. 3.9e-114;  
 Matches 246; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 24 DVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSSLPQSQRVYALHILTNVADETI 83  
 DB 1 DVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSSLPQSQRVYALHILTNVADETI 60

QY 84 SVAIDVTNVYIMGYRAGDTSYFFNGASATEAAKYVEKDMARKVTLPYSGNYERLQTAAGK 143  
 DB 61 SVAIDVTNVYIMGYRAGDTSYFFNEASATEAAKYVEKDMARKVTLPYSGNYERLQTAAGK 120

QY 144 IRENIPGLPALDSAITTLFYNNANSAASALMVLIOSTSEAAKYKFEIQOIGKRVKDTFL 203  
 DB 121 IRENIPGLPALDSAITTLFYNNANSAASALMVLIOSTSEAAKYKFEIQOIGKRVKDTFL 180

QY 204 PSIAIISLNSWSALSQIQIASTNNGQFESPVLINAQNRVTITNVDAAGVTSNTALL 263  
 DB 181 PSIAIISLNSWSALSQIQIASTNNGQFESPVLINAQNRVTITNVDAAGVTSNTALL 240

QY 264 LNRNMA 270  
 DB 241 LNRNMA 247

RESULT 4

US-10-375-209A-39  
 ; Sequence 39, Application US/10375209A  
 ; Publication No. US20030215421A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: McDonald, John R.  
 ; APPLICANT: Coggin, Philip  
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING SECONDARY TISSUE DAMAGE AN  
 ; TITLE OF INVENTION: OTHER INFLAMMATORY CONDITIONS AND DISORDERS  
 ; FILE REFERENCE: 25020-601E  
 ; CURRENT APPLICATION NUMBER: US/10/375,209A



; CURRENT FILING DATE: 2003-02-24  
 ; NUMBER OF SEQ ID NOS: 93  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 39  
 ; LENGTH: 247  
 ; TYPE: PRT  
 ; ORGANISM: Trichosantheus kirilowii  
 US-10-375-209A-39

Query Match 84.8%; Score 1214; DB 15; Length 247;  
 Best Local Similarity 99.6%; Pred. No. 3.9e-114;  
 Matches 246; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 24 DVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSSLPQSORYALIHLYTNYADETI 83  
 Db 1 DVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSSLPQSORYALIHLYTNYADETI 60  
 QY 84 SVAIDVTNVIYINGYRAGDTSYFFNGASATEAAKYVFKDAMRKVTLTPYSGNYERLQTAAGK 143  
 Db 61 SVAIDVTNVIYINGYRAGDTSYFFNEASATEAAKYVFKDAMRKVTLTPYSGNYERLQTAAGK 120  
 QY 144 IRENIPGLPALDSAITTLFYNNANSAASALMVLIQSTSEAAKYKFIQQIGKRVDTFL 203  
 Db 121 IRENIPGLPALDSAITTLFYNNANSAASALMVLIQSTSEAAKYKFIQQIGKRVDTFL 180  
 QY 204 PSLAIISLNSWSALSQKIQIASTNNGQFSPVVLINAGNQRTITNVDAGVVTSNIAL 263  
 Db 181 PSLAIISLNSWSALSQKIQIASTNNGQFSPVVLINAGNQRTITNVDAGVVTSNIAL 240  
 QY 264 LNRNMA 270  
 Db 241 LNRNMA 247

## RESULT 5

US-10-127-890-6  
 ; Sequence 6, Application US/10127890  
 ; Publication No. US20030166196A1  
 GENERAL INFORMATION:  
 APPLICANT: Better, Marc D.  
 Carroll, Stephen F.  
 Studnika, Gary M.  
 TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
 NUMBER OF SEQUENCES: 173  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
 STREET: 500 West Madison Street, 34th floor  
 CITY: Chicago  
 STATE: Illinois  
 COUNTRY: USA  
 ZIP: 60661  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/10/127,890  
 FILING DATE: 23-Apr-2002  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/08/646,360  
 FILING DATE: 13-MAY-1996  
 APPLICATION NUMBER: PCT/US94/05348  
 FILING DATE: 12-MAY-1994  
 APPLICATION NUMBER: US 08/064,691  
 FILING DATE: 12-MAY-1993  
 APPLICATION NUMBER: US 07/988,430  
 FILING DATE: 09-DEC-1992  
 APPLICATION NUMBER: US 07/901,707  
 FILING DATE: 19-JUN-1992  
 APPLICATION NUMBER: US 07/787,567

; FILING DATE: 04-NOV-1991  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: McNicholas, Janet M.  
 ; REGISTRATION NUMBER: 32,918  
 ; REFERENCE/DOCKET NUMBER: 200-70.P4  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 312/707-8889  
 ; TELEFAX: 312/707-9155  
 ; TELEX: 650 388-1248  
 ; INFORMATION FOR SEQ ID NO: 6:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 247 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 6:  
 US-10-127-890-6

Query Match 83.9%; Score 1202; DB 14; Length 247;  
 Best Local Similarity 98.8%; Pred. No. 6.4e-113;  
 Matches 244; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 24 DVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSSLPQSORYALIHLYTNYADETI 83  
 Db 1 DVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSSLPQSORYALIHLYTNYADETI 60  
 QY 84 SVAIDVTNVIYINGYRAGDTSYFFNGASATEAAKYVFKDAMRKVTLTPYSGNYERLQTAAGK 143  
 Db 61 SVAIDVTNVIYINGYRAGDTSYFFNEASATEAAKYVFKDAMRKVTLTPYSGNYERLQTAAGK 120  
 QY 144 IRENIPGLPALDSAITTLFYNNANSAASALMVLIQSTSEAAKYKFIQQIGKRVDTFL 203  
 Db 121 IRENIPGLPALDSAITTLFYNNANSAASALMVLIQSTSEAAKYKFIQQIGKRVDTFL 180  
 QY 204 PSLAIISLNSWSALSQKIQIASTNNGQFSPVVLINAGNQRTITNVDAGVVTSNIAL 263  
 Db 181 PSLAIISLNSWSALSQKIQIASTNNGQFSPVVLINAGNQRTITNVDAGVVTSNIAL 240  
 QY 264 LNRNMA 270  
 Db 241 LNRNMA 247

## RESULT 6

US-09-792-793A-34  
 ; Sequence 34, Application US/09792793A  
 ; Patent No. US20020168370A1  
 GENERAL INFORMATION:  
 APPLICANT: McDonald, John R.  
 Coggins, Philip  
 TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING SECONDARY TISSUE DAMAGE AND  
 OTHER INFLAMMATORY CONDITIONS AND DISORDERS  
 FILE REFERENCE: 25020-601D  
 CURRENT APPLICATION NUMBER: US/09/792,793A  
 CURRENT FILING DATE: 2001-02-22  
 NUMBER OF SEQ ID NOS: 93  
 SOFTWARE: PatentIn Ver. 2.0  
 SEQ ID NO 34  
 LENGTH: 247  
 TYPE: PRT  
 ORGANISM: Bryonia dioica  
 US-09-792-793A-34

Query Match 74.3%; Score 1064; DB 9; Length 247;  
 Best Local Similarity 85.8%; Pred. No. 5.7e-99;  
 Matches 212; Conservative 19; Mismatches 16; Indels 0; Gaps 0;  
 QY 24 DVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSSLPQSORYALIHLYTNYADETI 83  
 Db 1 DVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSSLPQSORYALIHLYTNYADETI 60  
 QY 84 SVAIDVTNVIYINGYRAGDTSYFFNGASATEAAKYVFKDAMRKVTLTPYSGNYERLQTAAGK 143

Db 61 SVAVDTNVTNYINGVYAGDVSYFFNEASATEAAKFVKDAKKVTLTPYSGNYERLQTAAGK 120  
QY 144 IRENIPGLPALDSAITTLFFYNNANSASALMVLQSTSEAAKYKFIHQIGKRVDTKTEL 203  
Db 121 IRENIPGLPALDSAITTLFFYNNANSASALMVLQSTSEAAKYKFIHQIGKRVDTKTEL 180  
QY 204 PSLAIIISLNSWSALSQKIQIASTNNQFESPVVLINAGNQRTITNVDAVVVTSNIAL 263  
Db 181 PSLAIIISLNSWSALSQKIQIASTNNQFESPVVLINAGNQRTITNVDAVVVTSNIAL 240  
QY 264 LNRNNMA 270  
Db 241 LNRNNIA 247  
RESULT 7  
US-10-375-209A-34  
; Sequence 34, Application US/10375209A  
; Publication No. US20030215421A1  
; GENERAL INFORMATION:  
; APPLICANT: McDonald, John R.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING SECONDARY TISSUE DAMAGE AND  
; TITLE OF INVENTION: OTHER INFLAMMATORY CONDITIONS AND DISORDERS  
; FILE REFERENCE: 25020-60LE  
; CURRENT APPLICATION NUMBER: US/10/375,209A  
; CURRENT FILING DATE: 2003-02-24  
; NUMBER OF SEQ ID NOS: 93  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 34  
; LENGTH: 247  
; TYPE: PRT  
; ORGANISM: Bryonia dioica  
US-10-375-209A-34  
Query Match 74.3%; Score 1064; DB 15; Length 247;  
Best Local Similarity 85.8%; Pred. No. 5.7e-99;  
Matches 212; Conservative 19; Mismatches 16; Indels 0; Gaps 0;  
QY 24 DVSPRLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSSLPQSGORYALIHLYTNYADETI 83  
Db 1 DVSPRLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSSLPQSGORYALIHLYTNYADETI 60  
QY 84 SVAVDTNVTNYINGVYAGDVSYFFNEASATEAAKFVKDAKKVTLTPYSGNYERLQTAAGK 143  
Db 61 SVAVDTNVTNYINGVYAGDVSYFFNEASATEAAKFVKDAKKVTLTPYSGNYERLQTAAGK 120  
QY 144 IRENIPGLPALDSAITTLFFYNNANSASALMVLQSTSEAAKYKFIHQIGKRVDTKTEL 203  
Db 121 IRENIPGLPALDSAITTLFFYNNANSASALMVLQSTSEAAKYKFIHQIGKRVDTKTEL 180  
QY 204 PSLAIIISLNSWSALSQKIQIASTNNQFESPVVLINAGNQRTITNVDAVVVTSNIAL 263  
Db 181 PSLAIIISLNSWSALSQKIQIASTNNQFESPVVLINAGNQRTITNVDAVVVTSNIAL 240  
QY 264 LNRNNMA 270  
Db 241 LNRNNIA 247  
RESULT 8  
US-10-127-890-7  
; Sequence 7, Application US/10127890  
; Publication No. US20030166196A1  
; GENERAL INFORMATION:  
; APPLICANT: Better, Marc D.  
; Carroll, Stephen F.  
; Studnika, Gary M.  
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
; Proteins  
; NUMBER OF SEQUENCES: 173  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.

STREET: 500 West Madison Street, 34th floor  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60661  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/127,890  
FILING DATE: 23-Apr-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/646,360  
FILING DATE: 13-MAY-1996  
APPLICATION NUMBER: PCT/US94/05348  
FILING DATE: 12-MAY-1994  
APPLICATION NUMBER: US 08/064,691  
FILING DATE: 12-MAY-1993  
APPLICATION NUMBER: US 07/988,430  
FILING DATE: 09-DEC-1992  
APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-JUN-1992  
APPLICATION NUMBER: US 07/787,567  
FILING DATE: 04-NOV-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: McNicholas, Janet M.  
REGISTRATION NUMBER: 32,918  
REFERENCE/DOCKET NUMBER: 200-70.P4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/707-9155  
TELEFAX: 312/707-8889  
TELEX: 650 388-1248  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 263 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 7:  
US-10-127-890-7  
Query Match 56.3%; Score 806.5; DB 14; Length 263;  
Best Local Similarity 63.5%; Pred. No. 6.9e-73;  
Matches 165; Conservative 38; Mismatches 56; Indels 1; Gaps 1;  
QY 24 DVSPRLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSSLPQSGORYALIHLYTNYADETI 83  
Db 1 DVSPRLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSSLPQSGORYALIHLYTNYADETI 60  
QY 84 SVAVDTNVTNYINGVYAGDVSYFFNEASATEAAKFVKDAKKVTLTPYSGNYERLQTAAGK 143  
Db 61 SVAVDTNVTNYINGVYAGDVSYFFNEASATEAAKFVKDAKKVTLTPYSGNYERLQTAAGK 120  
QY 144 IRENIPGLPALDSAITTLFFYNNANSASALMVLQSTSEAAKYKFIHQIGKRVDTKTEL 203  
Db 121 IRENIPGLPALDSAITTLFFYNNANSASALMVLQSTSEAAKYKFIHQIGKRVDTKTEL 180  
QY 204 PSLAIIISLNSWSALSQKIQIASTNNQFESPVVLINAGNQRTITNVDAVVVTSNIAL 263  
Db 181 PSLAIIISLNSWSALSQKIQIASTNNQFESPVVLINAGNQRTITNVDAVVVTSNIAL 240  
QY 264 LNRNNMA 270  
Db 241 LNRNNIA 247  
RESULT 9  
US-10-127-890-5  
; Sequence 5, Application US/10127890  
; Publication No. US20030166196A1

GENERAL INFORMATION:  
APPLICANT: Better, Marc D.  
Carroll, Stephen F.  
Studnika, Gary M.  
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating Proteins  
NUMBER OF SEQUENCES: 173  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
STREET: 500 West Madison Street, 34th floor  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60661  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/127,890  
FILING DATE: 23-Apr-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/646,360  
FILING DATE: 13-MAY-1996  
APPLICATION NUMBER: PCT/US94/05348  
FILING DATE: 12-MAY-1994  
APPLICATION NUMBER: US 08/064,691  
FILING DATE: 12-MAY-1993  
APPLICATION NUMBER: US 07/988,430  
FILING DATE: 09-DEC-1992  
APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-JUN-1992  
APPLICATION NUMBER: US 07/787,567  
FILING DATE: 04-NOV-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: McNicholas, Janet M.  
REGISTRATION NUMBER: 32,918  
REFERENCE/DOCKET NUMBER: 200-70.P4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/707-8889  
TELEFAX: 312/707-9155  
TELEX: 650 388-1248  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 248 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 5:  
US-10-127-890-5  
50.1%; Score 717.5; DB 14; Length 248;  
Best Local Similarity 58.6%; Pred. No. 6.3e-64;  
Matches 146; Conservative 49; Mismatches 51; Indels 3; Gaps 3;  
QY 24 DVSRFLSGATSSSYGVFISNLKALPNERKLYDIPLLRSSLPQSQRVALIHLTNVADETI 83  
DB 1 DVSRFLSGSSSTSYKIGDLRKALPNSGTVYNTILLSSASGASRYTLTSLNDGKAI 60  
QY 84 SVAIDVTNVIMGYRAGDTSYFFNGASATEAKVFKDAMRKVTLPYSGNYERLQTAAGK 143  
DB 61 TVAVDVSQLVIMGYLVNSTSYFFNESDAKLASQVYFKGS-TIVTLPSYSGNYEKLTAAKG 119  
QY 144 IRENIPGLPALDSAITTLTFYNNASASALMWLIQSTSEAKYKFTYEQIGKRVDTFL 203  
DB 120 IREKIPGLPALDSALTTFIHYDSTAAAPLVLTQTAASRKYTEGQIERISKNQV 179  
QY 204 PSIAISLENS-WSALSQIQIASTNNQFESPVLINAGNQRTVITNDAGVTSNIAL 262  
DB 180 PSIAISLENSLWSLSKQIQIACTNNGTKFTPVITDDKQQRVEINVTSKVVTNQL 239

QY 263 LLN-RUNMA 270  
DB 240 LLNKKQVA 248  
RESULT 10  
US-10-127-890-4  
Sequence 4, Application US/10127890  
Publication No. US20030166196A1  
GENERAL INFORMATION:  
APPLICANT: Better, Marc D.  
Carroll, Stephen F.  
Studnika, Gary M.  
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating Proteins  
NUMBER OF SEQUENCES: 173  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
STREET: 500 West Madison Street, 34th floor  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60661  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/127,890  
FILING DATE: 23-Apr-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/646,360  
FILING DATE: 13-MAY-1996  
APPLICATION NUMBER: PCT/US94/05348  
FILING DATE: 12-MAY-1994  
APPLICATION NUMBER: US 08/064,691  
FILING DATE: 12-MAY-1993  
APPLICATION NUMBER: US 07/988,430  
FILING DATE: 09-DEC-1992  
APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-JUN-1992  
APPLICATION NUMBER: US 07/787,567  
FILING DATE: 04-NOV-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: McNicholas, Janet M.  
REGISTRATION NUMBER: 32,918  
REFERENCE/DOCKET NUMBER: 200-70.P4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/707-8889  
TELEFAX: 312/707-9155  
TELEX: 650 388-1248  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 263 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-10-127-890-4  
49.1%; Score 703; DB 14; Length 263;  
Best Local Similarity 61.2%; Pred. No. 2e-62;  
Matches 148; Conservative 30; Mismatches 62; Indels 2; Gaps 2;  
QY 24 DVSRFLSGATSSSYGVFISNLKALPNERKLYDIPLLRSSLPQSQRVALIHLTNVADETI 83  
DB 1 DVNFDLSTATAKTYTKFIEDFRATLPFSHKVYDIPLLYSTISDSRRFILLDLTSYAVETI 60  
QY 84 SVAIDVTNVIMGYRAGDTSYFFNGASATEAKVFKDAMRKVTLPYSGNYERLQTAAGK 143  
DB 61 SVAIDVTNVIMGYRAGDTSYFFNGASATEAKVFKDAMRKVTLPYSGNYERLQTAAGK 118

APPLICANT: Better, Marc D.  
Carroll, Stephen F.  
Studnika, Gary M.  
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
Proteins  
NUMBER OF SEQUENCES: 173  
CORRESPONDENCE ADDRESS: ADDRESS: Held & Malloy, Ltd.  
STREET: 500 West Madison Street, 34th Floor  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60661  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/127,890  
FILING DATE: 23-Apr-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/646,360  
FILING DATE: 13-MAY-1996  
APPLICATION NUMBER: PCT/US94/05348  
FILING DATE: 12-MAY-1994  
APPLICATION NUMBER: US 08/064,691  
FILING DATE: 12-MAY-1993  
APPLICATION NUMBER: US 07/988,430  
FILING DATE: 09-DEC-1992  
APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-JUN-1992  
APPLICATION NUMBER: US 07/787,567  
FILING DATE: 04-NOV-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: McNicholas, Janet M.  
REGISTRATION NUMBER: 32,918  
REFERENCE/DOCKET NUMBER: 200-70.P4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/707-8889  
TELEFAX: 312/707-9155  
TELEX: 650 388-1248  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 267 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-10-127-890-1  
Query Match 29.3%; Score 419; DB 14; Length 267;  
Best Local Similarity 36.6%; Pred. No. 1.1e-33;  
Matches 93; Conservative 60; Mismatches 85; Indels 16; Gaps 7;  
QY 25 VSRFLSGATSSYGVFISNLRKALPNERKL-YDIPLL--RSSLPQSQRYALHILHTNYADE 81  
DB 9 INFTTAGATVQSYTFIRAVRGLTTGADVREHPVLPNVRGLPINQRFILVELSHAEEL 68  
QY 82 TISVAIDVTNVYIMVGRAGDTSYFF--NGASATEAAKYVFKDMRKVTLTPYSGNYERLOT 139  
DB 69 SVTLALDVTNAYVVGVRAGNSAYFFHPDQDEAEATHLFTDVQNRYYTFAGGNYDRLEQ 128  
QY 140 AAGKIRENIPGLPALDSAITTLFYNNAN-----SAASALMWLIQSTSEAAKYKFEQQI 194  
DB 129 LAGNLRNIELGNGLPEEAISALYYSTGTQTLPTLARSFIICIQMISEAARFYIEGEM 188  
QY 195 GKRV--DKTFPLPSLAISLNSWSALSQIQIASTNNQGFSPVVLINAGNQRTVINVD 252  
DB 189 RTRIRNRRSAPPSPVITLNSWGRSLTAIQ--ESNQGFASPIQLQRRNGSKFSV--YD 244  
QY 253 AGVVTNSIALLLNR 266

144 IRENIPGLPALDSAITTLFYNNANSAASALMWLIQSTSEAAKYKFEQQIGKRVDKTFL 203  
119 IRENIDGLPALSAITTLFYNAQAPSALLVLIQTAAARFKYIERHVAKYVATNFK 178  
204 PSALISLNSWSALSQIQIASTNNQGFSPVVLINAGNQRTVINVDAGVVTNSIALL 263  
179 PNLAIISLNSWSALSQIQIPLAQOGKFRNPVDLIKPTGERQVNTVDSVVYKGNIKLL 238  
264 LN 265  
239 LN 240  
RESULT 11  
US-10-282-935-1  
Sequence 1, Application US/10282935  
Publication No. US20030143193A1  
GENERAL INFORMATION:  
APPLICANT: VITETTA, ELLEN S.  
APPLICANT: GHETTIE, VICTOR F.  
APPLICANT: SMALLSHAW, JOAN  
APPLICANT: BALUNA, ROXANA G.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODIFYING TOXIC EFFECTS OF  
PROTEINACEOUS COMPOUNDS  
FILE REFERENCE: UTSD:884US  
CURRENT APPLICATION NUMBER: US/10/282,935  
CURRENT FILING DATE: 2002-10-29  
PRIOR APPLICATION NUMBER: 09/538,873  
PRIOR FILING DATE: 2000-03-30  
PRIOR APPLICATION NUMBER: 60/126,826  
PRIOR FILING DATE: 1999-03-30  
NUMBER OF SEQ ID NOS: 23  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 1  
LENGTH: 267  
TYPE: PRP  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Peptide  
US-10-282-935-1  
Query Match 29.3%; Score 419; DB 14; Length 267;  
Best Local Similarity 36.6%; Pred. No. 1.1e-33;  
Matches 93; Conservative 60; Mismatches 85; Indels 16; Gaps 7;  
QY 25 VSRFLSGATSSYGVFISNLRKALPNERKL-YDIPLL--RSSLPQSQRYALHILHTNYADE 81  
DB 9 INFTTAGATVQSYTFIRAVRGLTTGADVREHPVLPNVRGLPINQRFILVELSHAEEL 68  
QY 82 TISVAIDVTNVYIMVGRAGDTSYFF--NGASATEAAKYVFKDMRKVTLTPYSGNYERLOT 139  
DB 69 SVTLALDVTNAYVVGVRAGNSAYFFHPDQDEAEATHLFTDVQNRYYTFAGGNYDRLEQ 128  
QY 140 AAGKIRENIPGLPALDSAITTLFYNNAN-----SAASALMWLIQSTSEAAKYKFEQQI 194  
DB 129 LAGNLRNIELGNGLPEEAISALYYSTGTQTLPTLARSFIICIQMISEAARFYIEGEM 188  
QY 195 GKRV--DKTFPLPSLAISLNSWSALSQIQIASTNNQGFSPVVLINAGNQRTVINVD 252  
DB 189 RTRIRNRRSAPPSPVITLNSWGRSLTAIQ--ESNQGFASPIQLQRRNGSKFSV--YD 244  
QY 253 AGVVTNSIALLLNR 266  
245 VSILIPILALMYVR 258  
RESULT 12  
US-10-127-890-1  
Sequence 1, Application US/10127890  
Publication No. US20030166196A1  
GENERAL INFORMATION:

```

Db      245 VSLIPPIIALMVYR 258
::|||::|

RESULT 13
US-10-440-796-1
; Sequence 1, Application US/10440796
; Publication No. US20040009148A1
; GENERAL INFORMATION:
; APPLICANT: VITETTA, ELLEN S.
; APPLICANT: GHETIE, VICTOR F.
; APPLICANT: SMALLSHAW, JOAN
; APPLICANT: BALUNA, ROXANA G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR AMELIORATING VASCULAR LEAK
; TITLE OF INVENTION: SYNDROME (VLS)
; FILE REFERENCE: UTSD:603
; CURRENT APPLICATION NUMBER: US/10/440,796
; CURRENT FILING DATE: 2003-05-19
; PRIOR APPLICATION NUMBER: US/09/538,873
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/126,826
; PRIOR FILING DATE: 1999-03-30
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-440-796-1

Query Match          29.3%; Score 419; DB 15; Length 267;
Best Local Similarity 36.6%, Pred.No.1.le-33;
Matches 93; Conservative 60; Mismatches 85; Indels 16; Gaps 7;

QY    25 VSFRLSGATSSSYGVFFISNLRKALPNERKL-YDIPLL--RSSLPGSORYALIHLYVADE 81
Db    : :: || | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
9 INFTTAGATVQSNTNFIRAVRGRLTTCADYVHEIFVLPNRVGLPINQRFFILVELSNHAEL 68
QY    82 TISVAIDVTNYVMGYRAGDTSTVF--NGASATEAAKYVKFDAMRKVTLPLVPSGNYERLQT 139
Db    :: || | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
69 SVTALDVTNAYVYGRAGNSAYFFHPDNQEDAEATHLTDDQNRYTFAFGGNYDRLEQ 128
QY    140 AAGKIRENIPLGLPALDSALTTLFYNNAN----SAASALMWLIQSTSEAAKYKFIEQQI 194
Db    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
129 LAGNLRENIELGNGLPEALSIAWYYSTGTQPTLARSFIICIMTSEAARFQYIEGEM 188
QY    195 GKRV--DKTEFLPLSLAIISLENSALSQIQIASITNNGQPESPVLINACNQRTVINVD 252
Db    : :: || | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
189 RTRIRYNRRSAPDPSPVITLLENSWGRLSTAIO--ESNQGAFASPQLQRNGSKFSV.-YD 244
QY    253 AGVTSNIALLLNR 266
Db    :: || | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
245 VSLIPPIIALMVYR 258

RESULT 14
US-10-083-336A-1
; Sequence 1, Application US/10083336A
; Publication No. US20030181665A1
; GENERAL INFORMATION:
; APPLICANT: Olson, Mark A
; APPLICANT: Millard, Charles B
; APPLICANT: Byrnie, Michael P
; APPLICANT: Wannemacher, Robert W
; TITLE OF INVENTION: Rigiv Vaccine and Methods of Making and Using Thereof
; FILE REFERENCE: P67452US0 (RIID 01-58)
; CURRENT APPLICATION NUMBER: US/10/083,336A
; CURRENT FILING DATE: 2002-05-21
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1

```

us-09-905-247-1.rapb

Wed Sep 15 10:32:04 2004

Search completed: September 10, 2004, 14:12:01  
Job time : 139 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 13:50:56 ; Search time 41 Seconds  
(without alignments)  
678.033 Million cell updates/sec

Title: US-09-905-247-1

Perfect score: 1432

Sequence: 1 MRFVLVSLILTLFTTFA.....AAMDVDPMTQSGCGSYAL 289

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR 78:\*

2: PIR1:\*

3: PIR3:\*

4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1422	99.3	289	1	RLT2T
2	1377	96.2	289	2	rRNA N-glycosidase
3	1187	82.9	247	2	karasurin C
4	1182	82.5	247	2	karasurin - Mongol
5	859.5	60.0	286	1	karasurin-B - Tric
6	802.5	56.0	277	2	rRNA N-glycosidase
7	775	54.1	278	2	beta-luffin - smoo
8	756	52.8	286	2	rRNA N-glycosidase
9	753	52.6	286	2	rRNA N-glycosidase
10	673	47.0	250	2	luffin-b - smooth
11	663.5	46.3	245	2	rRNA N-glycosidase
12	419	29.3	576	1	ricin D precursor
13	394.5	27.5	564	1	agglutinin precurs
14	356.5	24.9	570	2	agglutinin I precu
15	349	24.4	251	2	abrin (clone 7.2)
16	343	24.0	562	2	abrin-c precursor
17	340	23.7	528	2	abrin-d precursor
18	337.5	23.6	527	2	abrin-a precursor
19	333	23.3	528	1	mistletoe lectin I
20	322	22.5	254	2	rRNA N-glycosidase
21	319	22.3	316	2	rRNA N-glycosidase
22	285.5	19.9	272	2	antiviral protein
23	271.5	19.0	294	2	ribosome-inactivat
24	265	18.5	313	2	rRNA N-glycosidase
25	237.5	16.6	261	2	rRNA N-glycosidase
26	215.5	15.0	310	2	TAP-29 anti-HIV pr
27	198.5	13.9	278	2	TAP-29 anti-HIV pr
28	181.5	12.7	289	2	TAP-29 anti-HIV pr
29	168	11.7	45	2	TAP-29 anti-HIV pr

## ALIGNMENTS

### RESULT 1

#### RLT2T

rRNA N-glycosidase (EC 3.2.2.22) alpha-trichosanthin precursor [validated] - Mongolian sr  
N/Alternate names: alpha-TCS; type I ribosome-inactivating protein  
C:Species: Trichosanthes kirilowii (Mongolian snake-gourd)  
C:Date: 30-Sep-1988 #sequence revision 26-Jan-1996 #text\_change 23-Mar-2001  
C:Accession: J05666; A36274; J01093; A36273; J00003  
R:Shaw, P.C.; Yung, M.H.; Zhu, R.H.; Ho, W.K.K.; Ng, T.B.; Yeung, H.W.  
Gene 97, 267-272, 1991  
A>Title: Cloning of trichosanthin cDNA and its expression in Escherichia coli.  
A:Reference number: J05666; MUID:91153657; PMID:1999291  
A:Accession: J05666  
A:Molecule type: mRNA  
A:Residues: 1-289 <SHA>  
A:Cross-references: GB:M34858; NID:g170536; PIDN:AAA34207.1; PID:g170537  
A:Experimental source: tuber  
R:Chow, T.P.; Feldman, R.A.; Lovett, M.; Piatak, M.  
J. Biol. Chem. 265, 8670-8674, 1990  
A>Title: Isolation and DNA sequence of a gene encoding alpha-trichosanthin, a type I ribo  
A:Reference number: A36274; MUID:90256790; PMID:2341400  
A:Accession: A36274  
A:Molecule type: DNA  
A:Residues: 1-233, 'T', 235-246, 'M', 248-289 <CHO>  
A:Cross-references: GB:J05434; NID:g170534; PIDN:AAA34206.1; PID:g170535  
R:Zheng, H.G.; Wang, B.; Shao, P.Z.; Yang, X.R.  
Acta Genet. Sin. 21, 42-51, 1994  
A>Title: Cloning and DNA sequencing of the gene encoding Trichosanthin.  
A:Reference number: J01093; MUID:94271613; PMID:8003348  
A:Accession: J01093  
A:Molecule type: DNA  
A:Residues: 1-72, 'V', 74-90, 'S', 92-233, 'T', 235-267, 'D', 269-289 <ZHE>  
A:Cross-references: GB:S70176; NID:G547148; PIDN:AAH31048.1; PID:G547149  
R:Collins, E.J.; Robertus, J.D.; LoPresti, M.; Stone, K.L.; Williams, K.R.; Wu, P.; Hwang  
J. Biol. Chem. 265, 8665-8669, 1990  
A>Title: Primary amino acid sequence of alpha-trichosanthin and molecular models for abri  
A:Reference number: A36273; MUID:90256789; PMID:2341399  
A:Accession: A36273  
A:Molecule type: protein  
A:Residues: 24-270 <COL>  
R:Wang, Y.; Qian, R.Q.; Gu, Z.W.; Jin, S.W.; Zhang, L.Q.; Xia, Z.X.; Tian, G.Y.; Ni, C.Z.  
Pure Appl. Chem. 58, 789-798, 1986  
A>Title: Scientific evaluation of Tian Hua Fen (THF): history, chemistry and application.  
A:Reference number: J00003  
A:Accession: J00003  
A:Molecule type: protein  
A:Residues: 24-56, 'L', 58-59, 'I', 61-71, 'I', 73-81, 85-86, 'L', 88-92, 'DAGLPRNAVL', 93-142, 'GL',  
R:Huang, Q.; Liu, S.; Tang, Y.; Jin, S.; Wang, Y.  
submitted to the Brookhaven Protein Data Bank, July 1994  
A:Reference number: A67091; PDB:1MRJ  
A:Contents: annotation: X-ray crystallography, 1.6 angstroms, with adenine, residues 24-;  
R:Huang, Q.; Liu, S.; Tang, Y.; Jin, S.; Wang, Y.

submitted to the Brookhaven Protein Data Bank, July 1994  
A:Reference number: A67092; PDB:IMRK  
A:Contents: annotation; X-ray crystallography, 1.6 angstroms, with formycin, residues 24-27  
R:Xiong, J.P.; Xia, Z.X.; Wang, Y.  
submitted to the Brookhaven Protein Data Bank, December 1994  
A:Reference number: A66711; PDB:1TCS  
A:Contents: annotation; X-ray crystallography, 1.7 angstroms, with NADPH, residues 24-27  
R:Xiong, J.P.; Xia, Z.X.; Wang, Y.  
Nat. Struct. Biol. 1, 695-700, 1994  
A:Title: Crystal structure of trichosanthin-NADPH complex at 1.7 Angstroms resolution  
A:Reference number: A58622; MUID:95360714; PMID:7634073  
A:Contents: annotation; X-ray crystallography, 1.7 angstroms  
C:Comment: Alpha-trichosanthin has been used to induce abortions.  
C:Genetics:  
A:Gene: tcs  
C:Function:  
A:Description: hydrolyzes the N-glycosidic bond of a specific adenosine in 28S rRNA then  
C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
F:1-23/Domain: signal sequence #status predicted <SIG>  
F:24-270/Product: trichosanthin alpha #status experimental <WAT>  
F:27-266/Domain: rRNA N-glycosidase homology <RNG>  
F:271-289/Domain: carboxyl-terminal propeptide #status predicted <CTP>  
F:93,183,186/Active site: Tyr, Glu, Arg #status predicted

Query Match 99.3%; Score 1422; DB 1; Length 289;  
Best Local Similarity 99.3%; Pred. No. 6.2e-108;  
Matches 287; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MIRFLVLSLLILTLFTTPAVEGDSVRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60  
DB 1 MIRFLVLSLLILTLFTTPAVEGDSVRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60

QY 61 RSSLPGSORYALIHLYTNVADETISVAIDVTNNYIMGYRAGDTSYFFNGASATEAAKYVFK 120  
DB 61 RSSLPGSORYALIHLYTNVADETISVAIDVTNNYIMGYRAGDTSYFFNGASATEAAKYVFK 120

QY 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSASALMVLIOQS 180  
DB 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSASALMVLIOQS 180

QY 181 TSEAAKYKFIHQKGVKDKTFLPSLAISLNSWSALSQIQIASTNNNGQFESPVLIN 240  
DB 181 TSEAAKYKFIHQKGVKDKTFLPSLAISLNSWSALSQIQIASTNNNGQFESPVLIN 240

QY 241 AQNRVTITNDVAGVTSNIALLLNRNNMAAMDDVPMTQSGCGSYAL 289  
DB 241 AQNRVTITNDVAGVTSNIALLLNRNNMAAMDDVPMTQSGCGSYAI 289

RESULT 2  
JC5606  
N:Contains: karasurin A  
C:Species: Trichosanthes kirilowii var. japonica  
C:Date: 23-Sep-1997 #sequence\_revision 23-Sep-1997 #text\_change 19-Jul-2002  
C:Accession: JC5606; JC5033  
R:Mizukami, H.; Iida, K.; Kondo, T.; Ogiwara, Y.  
Biol. Pharm. Bull. 20, 711-713, 1997  
A:Title: Cloning and bacterial expression of a gene encoding ribosome-inactivating prote  
A:Reference number: JC5606; MUID:97356562; PMID:9212998  
A:Accession: JC5606  
A:Molecule type: DNA  
A:Residues: 1-289 <MIZ>  
A:Cross-references: DDBJ:AB000666; NID:G2329830; PIDN:BA21786.1; PID:G2329831  
R:Kondo, T.; Mizukami, H.; Takeda, T.; Ogiwara, Y.  
Biol. Pharm. Bull. 19, 1485-1489, 1996  
A:Title: Amino acid sequences and ribosome-inactivating activities of karasurin-B and ka  
A:Reference number: JC5032; MUID:97108848; PMID:8951169  
A:Accession: JC5033  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 22-270 <KON>

C:Comment: This protein is a ribosome-inactivating protein and exhibit cytotoxic, abortif  
C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
F:22-270/Product: karasurin C #status predicted <MAC>  
F:24-270/Product: karasurin A #status predicted <MAA>  
F:27-266/Domain: rRNA N-glycosidase homology <RNG>

Query Match 96.2%; Score 1377; DB 2; Length 289;  
Best Local Similarity 95.8%; Pred. No. 2.8e-104;  
Matches 277; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 MIRFLVLSLLILTLFTTPAVEGDSVRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60  
DB 1 MIRFLVLSLLILTLFTTPAVEGDSVRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60

QY 61 RSSLPGSORYALIHLYTNVADETISVAIDVTNNYIMGYRAGDTSYFFNGASATEAAKYVFK 120  
DB 61 RSSLPGSORYALIHLYTNVADETISVAIDVTNNYIMGYRAGDTSYFFNGASATEAAKYVFK 120

QY 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSASALMVLIOQS 180  
DB 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSASALMVLIOQS 180

QY 181 TSEAAKYKFIHQKGVKDKTFLPSLAISLNSWSALSQIQIASTNNNGQFESPVLIN 240  
DB 181 TSEAAKYKFIHQKGVKDKTFLPSLAISLNSWSALSQIQIASTNNNGQFESPVLIN 240

QY 241 AQNRVTITNDVAGVTSNIALLLNRNNMAAMDDVPMTQSGCGSYAL 289  
DB 241 AQNRVTITNDVAGVTSNIALLLNRNNMAAMDDVPMTQSGCGSYAI 289

RESULT 3  
JU0393  
karasurin - Mongolian snake-gourd  
C:Species: Trichosanthes kirilowii (Mongolian snake-gourd)  
C:Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 12-Apr-1995  
C:Accession: JU0393; PS0163  
R:Toyokawa, S.; Takeda, T.; Kato, Y.; Wakabayashi, K.; Ogiwara, Y.  
Chem. Pharm. Bull. 39, 1244-1249, 1991  
A:Title: The complete amino acid sequence of an abortifacient protein, karasurin.  
A:Reference number: JU0393; MUID:92005921; PMID:1914000  
A:Accession: JU0393  
A:Molecule type: protein  
A:Residues: 1-247 <TOY>  
A:Note: a sequence which lacks Ala-247 is also shown in this publication  
C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
C:Keywords: abortifacient  
F:4-243/Domain: rRNA N-glycosidase homology <RNG>

Query Match 82.9%; Score 1187; DB 2; Length 247;  
Best Local Similarity 97.2%; Pred. No. 5.8e-89;  
Matches 240; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 24 DVSPRLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSSLPGSORYALIHLYTNVADETI 83  
DB 1 DVSPRLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSSLPGSORYALIHLYTNVADETI 60

QY 84 SVADVTNNYIMGYRAGDTSYFFNGASATEAAKYVFKDAMRKVTLPYSGNYERLQTAAGK 143  
DB 61 SVADVTNNYIMGYRAGDTSYFFNGASATEAAKYVFKDAMRKVTLPYSGNYERLQTAAGK 120

QY 144 IRENIPGLPALDSAITTLFYNNANSASALMVLIOQSTSEAAKYKFIHQKGVKDKTFL 203  
DB 121 IRENIPGLPALDSAITTLFYNNANSASALMVLIOQSTSEAAKYKFIHQKGVKDKTFL 180

QY 204 PSIAISLNSWSALSQIQIASTNNNGQFESPVLINNAQNRVTITNDVAGVTSNIAL 263  
DB 181 PSIAISLNSWSALSQIQIASTNNNGQFESPVLINNAQNRVTITNDVAGVTSNIAL 240

QY 264 LNENNMA 270  
DB 241 LNENNMA 247



## RESULT 4

JCS032  
 karasurin-B - Trichosanthes kirilowii var. japonica  
 C;Species: Trichosanthes kirilowii var. japonica  
 C;Date: 27-Feb-1997 #sequence\_revision 27-Feb-1997 #text\_change 23-May-1997  
 C;Accession: JCS032  
 R;Kondo, T.; Mizukami, H.; Takeda, T.; Ogihara, Y.  
 Biol. Pharm. Bull. 19, 1485-1489, 1996  
 A;Title: Amino acid sequences and ribosome-inactivating activities of karasurin-B and ka  
 A;Reference number: JCS032; MUID:97108848; PMID:8951169  
 A;Accession: JCS032  
 A;Status: preliminary  
 A;Molecule type: protein  
 A;Residues: 1-247 <KON>  
 C;Comment: This protein belongs to type I ribosomal-inactivating proteins which catalyti  
 C;Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
 F;4-243/Domain: rRNA N-glycosidase homology <RNG>

Query Match 82.5%; Score 1182; DB 2; Length 247;  
 Best Local Similarity 96.8%; Pred. No. 1.5e-88;  
 Matches 239; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 24 DVSFRLSGATSSSYGVFISNLRKALPNERKLDIPILRSLPGSQRYALIHLTNYADETI 93  
 Db 1 DVSFRLSGATSSSYGVFISNLRKALPNERKLDIPILRSLPGSQRYALIHLTNYADETI 60  
 QY 84 SVAIDVTNVIYNGYRAGDTSYFFNGASATEAAKYVFKDAMRKVTLPYSGNYERLQTAAGK 143  
 Db 61 SVAIDVTNVIYNGYRAGDTSYFFNEASATEAAKYVFKDARKVTLPYSGNYERLQTAAGK 120  
 QY 144 IRENTPLGLPALDSAITLTFYNNANSASALMVLIOSTSEAAKYFIEQIQKRVDTKFL 203  
 Db 121 IRENTPLGLPALDSAITLTFYNNANSASALMVLIOSTSEAAKYFIEQIQKRVDTKFL 180  
 QY 204 PSLAIISLNSWSALSLSKQIQAISTNNNGOPESPVMLINAQNRVTITNDVAGVTSNIAL 263  
 Db 181 PSLAIISLNSWSALSLSKQIQAISTNNNGOPETFPVVLINAQNRVTITNVNAGVTSNIAL 240

QY 264 LNRNMA 270

Db 241 LNRNMA 247

## RESULT 5

RPUUG  
 rRNA N-glycosidase (EC 3.2.2.22) alpha-momorcharin precursor [validated] - balsam pear  
 N;Alternate names: agglutinin; momordin-A; ribosome-inactivating protein momorcharin alp  
 C;Species: Momordica charantia (balsam pear, bitter gourd)  
 C;Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 15-Sep-2000  
 C;Accession: S14273; A61318; S16490; JN0628; S01670  
 R;Ho, W.K.K.; Liu, S.C.; Shaw, P.C.; Yeung, H.W.; Ng, T.B.; Chan, W.Y.  
 Biochim. Biophys. Acta 1088, 311-314, 1991  
 A;Title: Cloning of the cDNA of alpha-momorcharin: a ribosome inactivating protein.  
 A;Reference number: S14273; MUID:91159486; PMID:2001404

A;Accession: S14273

A;Molecule type: mRNA

A;Residues: 1-286 <HOW>

A;Cross-references: EMBL:X57682; NID:g19527; PID:CAA040869.1; PID:g19528

R;Li, S.S.L.

Experientia 36, 524-527, 1980

A;Title: Purification and partial characterization of two lectins from Momordica charant  
 A;Reference number: A61318; MUID:80201763; PMID:7379938  
 A;Accession: A61318

A;Molecule type: protein

A;Residues: 24-50 <LTA>

A;Note: as a lectin shows agglutinating activity for type-O red blood cells

R;Montecucchi, P.C.; Lazzerini, A.M.; Barbieri, L.; Stilpe, F.; Sorla, M.; Lappi, D.

Int. J. Pept. Protein Res. 33, 263-267, 1989

A;Title: N-terminal sequence of some ribosome-inactivating proteins.

A;Reference number: S16331; MUID:89326691; PMID:2753596

A;Accession: S16490

A;Molecule type: protein  
 A;Residues: 24-68 'X', 70 <MON>  
 R;Minami, Y.; Funatsu, G.

Biosci. Biotechnol. Biochem. 57, 1141-1144, 1993

A;Title: The complete amino acid sequence of momordin-a, a ribosome-inactivating protein  
 A;Reference number: JN0628; MUID:93372485; PMID:7763984  
 A;Accession: JN0628

A;Molecule type: protein

A;Residues: 24-107, 'Q', 109-123, 125-147, 'L', 149-154, 'I', 156-205, 'I', 207-208, 'L', 210-214, 'I'  
 A;Experimental source: seed

R;Ren, J.; Wang, Y.; Dong, Y.; Stuart, D.I.

submitted to the Brookhaven Protein Data Bank, January 1994

A;Reference number: A52272; PDB:1AHC

A;Contents: annotation; X-ray crystallography, 2.0 angstroms, residues 24-269

R;Husain, J.; Tickle, I.J.; Wood, S.P.

submitted to the Brookhaven Protein Data Bank, March 1994

A;Reference number: A52385; PDB:1MOM

A;Contents: annotation; X-ray crystallography, 2.16 angstroms, residues 24-86, 'L', 88-269

R;Huang, Q.; Liu, S.; Tang, Y.; Jin, S.; Wang, Y.

submitted to the Brookhaven Protein Data Bank, July 1994

A;Reference number: A67089; PDB:1MRH

A;Contents: annotation; X-ray crystallography, 2.0 angstroms, residues 24-77, 'R', 79-132, 'I'

C;Function:

A;Description: hydrolyzes the N-glycosidic bond of a specific adenosine in 28S rRNA there

C;Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology

C;Keywords: glycoprotein; glycosidase; hydrolase; lectin; seed; toxin

F;1-18/Domain: signal sequence #status predicted <SIG>

F;19-23/Domain: amino-terminal propeptide #status predicted <PRO>

F;24-269/Product: rRNA N-glycosidase alpha-momorcharin #status experimental <MAT>

F;27-266/Domain: rRNA N-glycosidase homology <RNG>

F;270-286/Domain: carboxyl-terminal propeptide #status predicted <CTP>

F;93.183.186/Active site: Tyr, Glu, Arg #status predicted

F;250/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 60.0%; Score 859.5; DB 1; Length 286;

Best Local Similarity 62.9%; Pred. No. 2.7e-62;

Matches 178; Conservative 40; Mismatches 64; Indels 1; Gaps 1;

QY 1 MIRFLVLSLLILTLFTTPAVEGDVSVFRLSGATSSSYGVFISNLRKALPNERKLDIPLL 60

Db 1 MSRFVLSFLILAIPLGGSIVKGDVSVFRLSGADPRSYGMFKDLRNLALPFREKVNIPL 60

QY 61 RSSLPGSORVALIHLTNVYADETISVAIDVTNVIYNGYRAGDTSYFFNGASATEAAKYVFK 120

Db 61 LPSVSGAGRYLMLHFNFDYDKTITVAVDVTNVIYNGYRAGDTSYFFNEPAELASQYVFR 120

QY 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITLTFYNNANSASALMVLIO 180

Db 121 DARRKITLPYSGNYERLQTAAGKPREKIPGLPALDSAITLHYDSTRAAGALLVLIQT 180

QY 181 TSEARVYKFTIQKRVDTKFLPSLAIISLNSWSALSLSKQIQAISTNNNGOPESPVVLIN 240

Db 181 TAEARVYKFTIQKRVDTKFLPSLAIISLNSWSALSLSKQIQAISTNNNGOPESPVVLIN 240

QY 241 AQNRVTITNVDAQVTSNTIALLRNNMAAMD-DYPMQTQSF 282

Db 241 NKGNRVQITNVTSKVTSNIQLLNTNRNIAEGONGDVSTTHGF 283

## RESULT 6

S22494  
 rRNA N-glycosidase (EC 3.2.2.22) alpha-luffin precursor - smooth loofah  
 N;Alternate names: protein synthesis inhibitor; ribosome-inactivating protein luffin-A  
 C;Species: Luffa cylindrica (smooth loofah)  
 C;Date: 12-Feb-1993 #sequence\_revision 12-Feb-1993 #text\_change 20-Aug-1999  
 C;Accession: S22494; S26390; JH0202; A32542  
 R;Kataoka, J.; Habuka, N.; Miyano, M.; Masuta, C.; Koiwai, A.  
 Plant Mol. Biol. 18, 1199-1202, 1992  
 A;Title: Nucleotide sequence of cDNA encoding alpha-luffin, a ribosome-inactivating prote  
 A;Reference number: S22494; MUID:92288316; PMID:1600156  
 A;Accession: S22494  
 A;Molecule type: mRNA  
 A;Residues: 1-277 <KAT>

A:Cross-references: EMBL:X62372; NID:g19149; PIDN:CAA44230.1; PID:g19150  
C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
F:26-264/Domain: rRNA N-glycosidase homology <RNG>

Query Match 54.1%; Score 775; DB 2; Length 278;  
Best Local Similarity 58.6%; Pred. No. 1.9e-55;  
Matches 163; Conservative 47; Mismatches 64; Indels 4; Gaps 3;

QY 1 MURFLVLSLLIULTFLTTPAVEG-DVSFRLSGATSSYGVFIINLRKALPNERKLYDIPL 59  
DB 1 MURFTFLSLLIILIIIAFFT--VEGANVSFSLSGADSKSYKFITALRKALPSKEKVSNIPL 58  
QY 60 LRSSLPGSORYALIHUHTNVADETIISVAIDVTNYVIMGYRAGDTSYFENGASATEAAKVF 119  
DB 59 LPLPSAGSAGRIILMQLSNYDAKAITMAIDVTNYVIMGYLVNSTSYFENGSDAKLASQV 118  
QY 120 KDAWRKVTLPYSGNVERLOTPAACKIRENIPGLPALDSAITTLTFYFNANGAASALMVLIQ 179  
DB 119 KGS-TIVTLTPYSGNYEELQNAAGKVRKEIPLGFRAFDTSITSLFHYDSTAAGAFVLIIQ 177  
QY 180 STSEAAARYKFEIOQIGKRVDKTFLPSLAITISLNSWSALSQIQIASTNNQGPSPVLI 239  
DB 178 TTAASRFKVEIQIIRIPKNEVPSPALSLNENWSALSQIQIQAQTNNGAFRTPVII 237  
QY 240 NAQQRVTITNDAGVVTNSIALLLNENNAAMDDVDP 277  
DB 238 DNKGQRVEIKDVNSKVVTNNIKLLNKONTAAAFDDGIP 275  
RESULT 8  
rRNA N-glycosidase (BC 3.2.2.22) momordin II - balsam apple  
C:Species: Momordica balsamina (balsam apple)  
C:Date: 25-Feb-1994 #sequence\_revision 03-Aug-1995 #text\_change 20-Aug-1999  
C:Accession: S25560  
R:Ortigao, M.; Better, M.  
Nucleic Acids Res. 20, 4662, 1992  
A:Title: Momordin II, a ribosome inactivating protein from Momordica balsamina  
A:Reference number: S25560; MUID:9302170; PMID:1408771  
A:Accession: S25560  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-286 <ORT>  
A:Cross-references: EMBL:Z12175; NID:g19525; PIDN:CAA78166.1; PID:g19526  
C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
C:Keywords: glycosidase; hydrolase  
F:27-264/Domain: rRNA N-glycosidase homology <RNG>

Query Match 52.8%; Score 756; DB 2; Length 286;  
Best Local Similarity 59.2%; Pred. No. 6.8e-54;  
Matches 157; Conservative 37; Mismatches 69; Indels 2; Gaps 2;

QY 1 MIRFLVLSLLIULTFLTTPAVEGDVSFRLSGATSSYGVFIINLRKALPNERKLYDIPL 60  
DB 1 MVKCLLSFLIIAIFIGVPTAKGDVNFDLSTATAKTYTKFIEDFRATLFPCHKVYDIPL 60  
QY 61 RSSLPGSORYALIHUHTNVADETIISVAIDVTNYVIMGYRAGDTSYFENGASATEAAKVF 120  
DB 61 YSTISDSRRFILLDTSYAYETIISVAIDVTNYVIMGYRTRDVSYPFK-ESPPEAYNILFK 119  
QY 121 DAWRKVTLPYSGNVERLOTPAACKIRENIPGLPALDSAITTLTFYFNANGAASALMVLIQ 180  
DB 120 -GTRKITLPTGYNENLQTPAAKIRENIDLGLPALDSAITTLTFYFNAQAPSALLVLIQ 178  
QY 181 TSEAAARYKFEIOQIGKRVDKTFLPSLAITISLNSWSALSQIQIASTNNQGPSPVLI 240  
DB 179 TAAARFKYIERHVAKYVATNEKPNLAIISLENQWSALSQIQIFLAQOQKGFNPVDLIK 238  
QY 241 AQQRVTITNDAGVVTNSIALLLN 265  
DB 239 PTGERFQVTNDSVVYKGNIKLLN 263







us-09-905-247-1.rpr

Wed Sep 15 10:32:04 2004

Search completed: September 10, 2004, 14:01:55  
Job time : 42 secs

GenCore version 5.1.6  
 Copyright (c) 1993 - 2004 CompuGen Ltd.  
 OM protein - protein search, using sw model  
 Run on: September 10, 2004, 13:47:56 ; Search time 118 Seconds  
 (without alignments)  
 772.752 Million cell updates/sec

Title: US-09-905-247-1  
 Perfect score: 1432  
 Sequence: 1 MIRFLVLSLLILFLTTTPA.....AAMDVDPMTQSGGCGSYAL 289

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues  
 Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : SPTREMBL.25.\*  
 1: sp\_archaea.\*  
 2: sp\_bacteria.\*  
 3: sp\_fungi.\*  
 4: sp\_human.\*  
 5: sp\_invertebrate.\*  
 6: sp\_mammal.\*  
 7: sp\_mbc.\*  
 8: sp\_organelle.\*  
 9: sp\_phage.\*  
 10: sp\_plant.\*  
 11: sp\_rodent.\*  
 12: sp\_virus.\*  
 13: sp\_vertebrate.\*  
 14: sp\_unclassified.\*  
 15: sp\_rvirus.\*  
 16: sp\_bacteriap.\*  
 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB ID	Description	SUMMARIES			
						1	2	3	4
1	1419	99.1	289	10	Q84SV8				
2	1418	99.0	289	10	Q94KE4				
3	1408	98.3	289	10	Q41216				
4	1262	88.1	270	10	Q41611				
5	1253	87.5	270	10	Q8LPU7				
6	1180	82.4	247	10	Q9LRE3				
7	775	54.1	278	10	Q00980				
8	758	52.9	286	10	Q9FUV7				
9	753	52.6	286	10	Q41257				
10	699	48.8	264	10	Q9FSH2				
11	469.5	32.8	136	10	Q8SAD7				
12	437.5	30.6	563	10	Q8GT32				
13	433.5	30.3	563	10	Q945S2				
14	430.5	30.1	563	10	Q04367				
15	423	29.5	541	10	Q41174				
16	414.5	28.9	564	10	Q9AVR2				

17	411.5	28.7	275	10	Q8H1Y4	Q8H1Y4 gynostemma
18	408.5	28.5	136	10	Q8SAG0	Q8SAG0 benincasa h
19	407	28.4	580	10	Q94BW4	Q94BW4 cinnamomum
20	403	28.1	580	10	Q94BW3	Q94BW3 cinnamomum
21	400.5	28.0	275	10	Q84LJ1	Q84LJ1 gynostemma
22	400	27.9	581	10	Q94BW5	Q94BW5 cinnamomum
23	396.5	27.7	277	10	Q8GV03	Q8GV03 gynostemma
24	396.5	27.7	549	10	Q9FV22	Q9FV22 cinnamomum
25	394.5	27.5	277	10	Q84JUR1	Q84JUR1 gynostemma
26	393.5	27.5	277	10	Q8GV11	Q8GV11 gynostemma
27	388.5	27.1	277	10	Q8H1Y5	Q8H1Y5 gynostemma
28	386.5	27.0	136	10	Q84LJ0	Q84LJ0 cucurbita m
29	385.5	26.9	136	10	Q84LJ9	Q84LJ9 gynostemma
30	381.5	26.6	277	10	Q8GV10	Q84LJ9 cucurbita m
31	377.5	26.4	136	10	Q8S2R5	Q8GV10 gynostemma
32	377.5	26.4	136	10	Q84LI8	Q8S2R5 cucurbita m
33	375.5	26.2	565	10	Q04071	Q84LI8 cucurbita m
34	362	25.3	566	10	Q04072	Q04071 sambucus ni
35	360.5	25.2	203	10	Q8RY69	Q04072 sambucus ni
36	357.5	25.0	570	10	Q22415	Q8RY69 gynostemma
37	356.5	24.9	570	10	Q41358	Q22415 sambucus ni
38	353.5	24.7	569	10	Q93543	Q41358 sambucus ni
39	349	24.4	252	10	Q38760	P93543 sambucus ni
40	347.5	24.3	293	10	Q8VYU0	Q38760 abrus preca
41	342	23.9	604	10	Q9M654	Q8VYU0 jatropa cu
42	339	23.7	565	10	Q8W243	Q9M654 polygonatum
43	338.5	23.6	293	10	Q8S452	Q8W243 viscum albu
44	338	23.6	248	10	Q7Y1U7	Q8S452 jatropa cu
45	336	23.5	248	10	Q7Y1U5	Q7Y1U7 gynostemma
						Q7Y1U5 gynostemma

## ALIGNMENTS

RESULT 1	
Q84SV8	PRELIMINARY; PRT; 289 AA.
ID	Q84SV8
AC	Q84SV8; Q84SV8
DT	01-JUN-2003 (Tremblrel. 24, Created)
DT	01-JUN-2003 (Tremblrel. 24, Last sequence update)
DT	01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE	Trichosanthin.
GN	TCS.
OS	Trichosanthes kirilowii (Mongolian snake-gourd).
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC	eurosid1; Cucurbitales; Cucurbitaceae; Trichosanthes.
OX	NCBI_TaxID=3677;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	Yuan H., Wang L., Wang Y., Liu T., An C., Chen Z.;
RT	"Trichosanthes kirilowii trichosanthin precursor (TCS) gene.";
RL	Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AY082349; AA072728.1; -
DR	GO; GO:0017149; F:protein biosynthesis inhibitor activity; IEA.
DR	GO; GO:0015070; F:toxin activity; IEA.
DR	GO; GO:0017448; F:negative regulation of protein biosynthesis; IEA.
DR	InterPro; IPR001574; RIP.
DR	Pfam; PF00161; RIP; 1.
DR	PRINTS; PRO0396; SHIGARICIN.
DR	PROSITE; PS00275; SHIGA_RICIN; 1.
SQ	SEQUENCE 289 AA; 31690 MW; B403148E968861FA CRC64;
Query Match	
Best Local Similarity 99.1%; Score 1419; DB 10; Length 289;	
Matches 286; Conservative 99.0%; 2; Mismatches 1; Indels 0; Gaps 0;	
QY	1 MIRFLVLSLLILFLTTTPAVEGDVSRFGATSSSYGVFISNLRKALPNERKIYDIPLL 60
Db	1 MIRFLVLSLLILFLTTTPAVEGDVSRFGATSSSYGVFISNLRKALPNERKIYDIPLL 60
QY	61 RSSLPQSQRVALHLYNADETISVALDVTNVMYMGYRAGDTSYFFNGASATSAAYVFK 120

```

Db 61 RSTLPGSQRYALHLNTYADETISVAIDVTNVYIMGYRAGDTSYFFNEASATEAAKYVFK 120
QY 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMVLQIS 180
Db 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMVLQIS 180
QY 181 TSEAARYKFTEQOIGKRVKDTFLPSLAISLENSWSALSQIOIASTNNQFSPVVLIN 240
Db 181 TSEAARYKFTEQOIGKRVKDTFLPSLAISLENSWSALSQIOIASTNNQFSPVVLIN 240
QY 241 AQNRVTITNDAGVVTNSIALLLNNRNMAAMDDDDVPMTQSFQCGSYAI 289
Db 241 AQNRVTITNDAGVVTNSIALLLNNRNMAAMDDDDVPMTQSFQCGSYAI 289

RESULT 2
Q94KE4 PRELIMINARY; PRT; 289 AA.
AC Q94KE4 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 25, Last annotation update)
DE Trichosanthin precursor (EC 3.2.2.22) (rRNA N-glycosidase).
GN TCS.
OS Trichosanthes kirilowii (Mongolian snake-gourd).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.
OX NCBI_taxid=3677;
RN [1]
SEQUENCE FROM N.A.
RA Yuan H., Wang L., Wang Y., An C., Chen Z.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC SPECIFIC ADENOSINE ON THE 28S RRNA.
CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
DR EMBL; AF367252; AAKS2960.1; -.
DR FIRM; J05032; J05032.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0017149; F:protein biosynthesis inhibitor activity; IEA.
DR GO; GO:0030598; F:rRNA N-glycosylase activity; IEA.
DR GO; GO:0015070; F:toxin activity; IEA.
DR GO; GO:0017148; P:negative regulation of protein biosynthesis; IEA.
DR InterPro; IPR001574; RIP.
DR Pfam; PF00161; RIP; 1.
DR PRINTS; PR00396; SHIGARICIN.
DR PROSITE; PS00275; SHIGA_RICIN; 1.
KW Hydrolase; Signal; Toxin.
FT SIGNAL 1 23
FT CHAIN 24 270
FT SEQUENCE 289 AA; 31706 MW; A6D5602549CA5657 CRC64;
QY 1 MIRFLVLSLLILTLFTTPAVEGDVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60
Db 1 MIRFLVLSLLILTLFTTPAVEGDVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60
QY 61 RSLPGSQRYALHLNTYADETISVAIDVTNVYIMGYRAGDTSYFFNEASATEAAKYVFK 120
Db 61 RSLPGSQRYALHLNTYADETISVAIDVTNVYIMGYRAGDTSYFFNEASATEAAKYVFK 120
QY 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMVLQIS 180
Db 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMVLQIS 180
QY 181 TSEAARYKFTEQOIGKRVKDTFLPSLAISLENSWSALSQIOIASTNNQFSPVVLIN 240
Db 181 TSEAARYKFTEQOIGKRVKDTFLPSLAISLENSWSALSQIOIASTNNQFSPVVLIN 240
QY 241 AQNRVTITNDAGVVTNSIALLLNNRNMAAMDDDDVPMTQSFQCGSYAI 289
Db 241 AQNRVTITNDAGVVTNSIALLLNNRNMAAMDDDDVPMTQSFQCGSYAI 289

Query Match 99.0%; Score 1418; DB 10; Length 289;
Best Local Similarity 98.6%; Pred. No. 4.2e-104;
Matches 285; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 MIRFLVLSLLILTLFTTPAVEGDVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60
Db 1 MIRFLVLSLLILTLFTTPAVEGDVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60
QY 61 RSLPGSQRYALHLNTYADETISVAIDVTNVYIMGYRAGDTSYFFNEASATEAAKYVFK 120
Db 61 RSLPGSQRYALHLNTYADETISVAIDVTNVYIMGYRAGDTSYFFNEASATEAAKYVFK 120
QY 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMVLQIS 180
Db 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMVLQIS 180
QY 181 TSEAARYKFTEQOIGKRVKDTFLPSLAISLENSWSALSQIOIASTNNQFSPVVLIN 240
Db 181 TSEAARYKFTEQOIGKRVKDTFLPSLAISLENSWSALSQIOIASTNNQFSPVVLIN 240
QY 241 AQNRVTITNDAGVVTNSIALLLNNRNMAAMDDDDVPMTQSFQCGSYAI 289
Db 241 AQNRVTITNDAGVVTNSIALLLNNRNMAAMDDDDVPMTQSFQCGSYAI 289

```

```

Db 241 AQNRVTITNDAGVVTNSIALLLNNRNMAAMDDDDVPMTQSFQCGSYAI 289
QY 241 AQNRVTITNDAGVVTNSIALLLNNRNMAAMDDDDVPMTQSFQCGSYAI 289

RESULT 3
Q41216 PRELIMINARY; PRT; 289 AA.
AC Q41216 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1996 (TrEMBLrel. 25, Last annotation update)
DE Trichosanthin (EC 3.2.2.22) (rRNA N-glycosidase).
GN TRICHOSANTHIN, TCS.
OS Trichosanthes kirilowii (Mongolian snake-gourd).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.
OX NCBI_taxid=3677;
RN [1]
SEQUENCE FROM N.A.
RA MEDLINE=94271613; PubMed=8003348;
RA Zheng H., Wang B., Shaw P., Yeung H.;
RT "[Cloning and DNA sequencing of the gene encoding trichosanthin].";
RL I Chuan Hsueh Pao 21:42-51(1994).
CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC SPECIFIC ADENOSINE ON THE 28S RRNA.
CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
DR EMBL; S70176; AAB31048.1; -.
DR PIR; JC5032; JC5032.
DR HSSP; P09989; IMRJ.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0017149; F:protein biosynthesis inhibitor activity; IEA.
DR GO; GO:0030598; F:rRNA N-glycosylase activity; IEA.
DR GO; GO:0015070; F:toxin activity; IEA.
DR GO; GO:0017148; P:negative regulation of protein biosynthesis; IEA.
DR InterPro; IPR001574; RIP.
DR Pfam; PF00161; RIP; 1.
DR PRINTS; PR00396; SHIGARICIN.
DR PROSITE; PS00275; SHIGA_RICIN; 1.
KW Hydrolase; Toxin.
FT SEQUENCE 289 AA; 31650 MW; 286AC14D48BCA175 CRC64;
QY 1 MIRFLVLSLLILTLFTTPAVEGDVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60
Db 1 MIRFLVLSLLILTLFTTPAVEGDVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60
QY 61 RSLPGSQRYALHLNTYADETISVAIDVTNVYIMGYRAGDTSYFFNEASATEAAKYVFK 120
Db 61 RSLPGSQRYALHLNTYADETISVAIDVTNVYIMGYRAGDTSYFFNEASATEAAKYVFK 120
QY 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMVLQIS 180
Db 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMVLQIS 180
QY 181 TSEAARYKFTEQOIGKRVKDTFLPSLAISLENSWSALSQIOIASTNNQFSPVVLIN 240
Db 181 TSEAARYKFTEQOIGKRVKDTFLPSLAISLENSWSALSQIOIASTNNQFSPVVLIN 240
QY 241 AQNRVTITNDAGVVTNSIALLLNNRNMAAMDDDDVPMTQSFQCGSYAI 289
Db 241 AQNRVTITNDAGVVTNSIALLLNNRNMAAMDDDDVPMTQSFQCGSYAI 289

Query Match 98.3%; Score 1408; DB 10; Length 289;
Best Local Similarity 97.9%; Pred. No. 2.6e-103;
Matches 283; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
QY 1 MIRFLVLSLLILTLFTTPAVEGDVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60
Db 1 MIRFLVLSLLILTLFTTPAVEGDVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60
QY 61 RSLPGSQRYALHLNTYADETISVAIDVTNVYIMGYRAGDTSYFFNEASATEAAKYVFK 120
Db 61 RSLPGSQRYALHLNTYADETISVAIDVTNVYIMGYRAGDTSYFFNEASATEAAKYVFK 120
QY 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMVLQIS 180
Db 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMVLQIS 180
QY 181 TSEAARYKFTEQOIGKRVKDTFLPSLAISLENSWSALSQIOIASTNNQFSPVVLIN 240
Db 181 TSEAARYKFTEQOIGKRVKDTFLPSLAISLENSWSALSQIOIASTNNQFSPVVLIN 240
QY 241 AQNRVTITNDAGVVTNSIALLLNNRNMAAMDDDDVPMTQSFQCGSYAI 289
Db 241 AQNRVTITNDAGVVTNSIALLLNNRNMAAMDDDDVPMTQSFQCGSYAI 289

RESULT 4
Q41611 PRELIMINARY; PRT; 270 AA.
AC Q41611 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1996 (TrEMBLrel. 25, Last annotation update)

```



DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Trichosanthes (EC 3.2.2.22) (rRNA N-glycosidase) (Fragment).  
 OS Trichosanthes kirilowii (Mongolian snake-gourd).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.  
 NCBI\_TaxID=3677;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Bao Y., Chu R., Han J., Zhang H., Pan N., Gu X., Chen Z.;  
 RT "Cloning and sequencing of trichosanthin gene and its expression in  
 RT Escherichia coli and tobacco plant.";  
 RL Sci. China B. 36:669-676 (0).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Xu L.;  
 RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.  
 CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 CC SPECIFIC ADENOSINE ON THE 28S RRNA.  
 CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 DR EMBL; U25675; AAA70036.1; -.  
 DR HSP; P09989; IMRJ.  
 DR GO; GO:0016787; F:hydrolase activity; IEA.  
 DR GO; GO:0017149; F:protein biosynthesis inhibitor activity; IEA.  
 DR GO; GO:0030598; F:rRNA N-glycosylase activity; IEA.  
 DR GO; GO:0015070; F:toxin activity; IEA.  
 DR GO; GO:0017148; P:negative regulation of protein biosynthesis; IEA.  
 DR InterPro; IPR001574; RIP.  
 DR Pfam; PF00161; RIP; 1.  
 DR PRINTS; PR00396; SHIGARICIN  
 DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
 DR Hydrolase; Toxin.  
 FT NON\_TER 270 270  
 SQ SEQUENCE 270 AA; 29993 MW; 3D73FB461EA8B8D4 CRC64;  
 Query Match 89.1%; Score 1262; DB 10; Length 270;  
 Best Local Similarity 94.8%; Pred. No. 8.4e-92;  
 Matches 256; Conservative 6; Mismatches 8; Indels 0; Gaps 0;  
 QY 1 MIRFLVLSLLILTLFTTTPAVEGDVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60  
 Db 1 MIRFLVLSLLILTLFTTTPAVEGDVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60  
 QY 61 RSLPGSORYALIHNTVYADETISVAIDVNTVYIMGYRAGDTSYFFNGASATEAAKYVFK 120  
 Db 61 RSTLPGCERYALIHNTVYADETISVAIDVNTVYIMGYRAGDTFFYFNEASATEAAKYVFK 120  
 QY 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMVLIQS 180  
 Db 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMVLIQS 180  
 QY 181 TSEAAKYKFIEQQIGKRVDTKFLPSLAISLNSWSALSQIQIASTNNGQFESPVLIN 240  
 Db 181 TSEAAKYKFIEQQIGKRVDTKFLPSLAISLNSWSALSQIQIASTNNGQFESPVLIN 240  
 QY 241 AQNRVTITNDAGVVTNSIALLLNRNMA 270  
 Db 241 AQNRVTITNDAGVVTNSIALLLNRNMA 270  
 RESULT 5  
 Q8LPV7 ID Q8LPV7 PRELIMINARY; PRT; 270 AA.  
 AC Q8LPV7;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Trichosanthin precursor (EC 3.2.2.22) (rRNA N-glycosidase) (Fragment).  
 GN TCS.  
 OS Trichosanthes kirilowii (Mongolian snake-gourd).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OC eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.  
 NCBI\_TaxID=3677;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Yuan H., Wang L., Wang Y., Liu T., An C., Chen Z.;  
 RT "Trichosanthes kirilowii trichosanthin precursor (TCS) gene.";  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 CC SPECIFIC ADENOSINE ON THE 28S RRNA.  
 CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 DR EMBL; AY082348; AAM22782.1; -.  
 DR PIR; JCS032; JCS032.  
 DR GO; GO:0016787; F:hydrolase activity; IEA.  
 DR GO; GO:0017149; F:protein biosynthesis inhibitor activity; IEA.  
 DR GO; GO:0030598; F:rRNA N-glycosylase activity; IEA.  
 DR GO; GO:0015070; F:toxin activity; IEA.  
 DR GO; GO:0017148; P:negative regulation of protein biosynthesis; IEA.  
 DR InterPro; IPR001574; RIP.  
 DR Pfam; PF00161; RIP; 1.  
 DR PRINTS; PR00396; SHIGARICIN  
 DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
 DR Hydrolase; Signal; Toxin.  
 KW SIGNAL 1 23  
 FT CHAIN 24 >270  
 FT NON\_TER 270 270  
 SQ SEQUENCE 270 AA; 29683 MW; 531713B754F9B769 CRC64;  
 Query Match 87.5%; Score 1253; DB 10; Length 270;  
 Best Local Similarity 94.8%; Pred. No. 4.3e-91;  
 Matches 256; Conservative 5; Mismatches 9; Indels 0; Gaps 0;  
 QY 1 MIRFLVLSLLILTLFTTTPAVEGDVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60  
 Db 1 MIRFLVLSLLILTLFTTTPAVEGDVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60  
 QY 61 RSLPGSORYALIHNTVYADETISVAIDVNTVYIMGYRAGDTSYFFNGASATEAAKYVFK 120  
 Db 61 RSLSGSORYALIHNTVYADETISVAIDVNTVYIMGYRAGDISYFFNEASATEAAKYVFK 120  
 QY 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMVLIQS 180  
 Db 121 DAKRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMVLIQS 180  
 QY 181 TSEAAKYKFIEQQIGKRVDTKFLPSLAISLNSWSALSQIQIASTNNGQFESPVLIN 240  
 Db 181 TSEAAKYKFIEQQIGKRVDTKFLPSLAISLNSWSALSQIQIASTNNGQFESPVLIN 240  
 QY 241 AQNRVTITNDAGVVTNSIALLLNRNMA 270  
 Db 241 AQNRVTITNDAGVVTNSIALLLNRNMA 270  
 RESULT 6  
 Q9LRE3 ID Q9LRE3 PRELIMINARY; PRT; 247 AA.  
 AC Q9LRE3;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Trichobakin (EC 3.2.2.22) (rRNA N-glycosidase) (Fragment).  
 GN TBK.  
 OS Trichosanthes sp. Bac Kan 8-98.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.  
 NCBI\_TaxID=118182;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Phan Van Chi, Nguyen Thuy Ha, Le Tran Binh;  
 RT "Genomic DNA Clone for mature typ-1 ribosome-inactivating protein from  
 RT Trichosanthes sp. sample 01 Bac Kan 8-98 Vien CNSH (Hanoi).";  
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE

CC SPECIFIC ADENOSINE ON THE 28S RNA.  
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 DR EMBL; AB039324; BAA92530.1; -.  
 DR HSP; P09989; 1MRJ.  
 DR GO; GO:0016787; F:hydrolase activity; IEA.  
 DR GO; GO:0017149; F:protein biosynthesis inhibitor activity; IEA.  
 DR GO; GO:0030598; F:RNA N-glycosylase activity; IEA.  
 DR GO; GO:0015070; F:toxin activity; IEA.  
 DR GO; GO:0017148; P:negative regulation of protein biosynthesis; IEA.  
 DR InterPro; IPR001574; RIP.  
 DR Pfam; PF00161; RIP; 1.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
 KW Hydrolase; Toxin.  
 FT NON TER 247 247  
 SQ SEQUENCE 247 AA; 27199 MW; 89811AC32892F03F CRC64;  
 Query Match 82.4%; Score 1180; DB 10; Length 247;  
 Best Local Similarity 96.8%; Pred. No. 2.3e-85;  
 Matches 239; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
 QY 24 DVSRFLSGATSSSYGVFISNLRKALPNERKLYDIPLLESSLPGSQRYALIHLTNYADETI 83  
 DB 1 DVSRFLSGATSSSYGVFISNLRKALPNERKLYDIPLLESSLPGSQRYALIHLTNYADETI 60  
 QY 84 SVAIDVTNVYMGYRAGDTSYFFNGASATEAAKYFKDAMRKVTLPSYGNVERLQTAAGK 143  
 DB 61 SVAIDVTNVYMGYRAGDTSYFFNGASATEAAKYFKDAMRKVTLPSYGNVERLQTAAGK 120  
 QY 144 IRENPLGLPALDSAITTLFFYNANSAASALMVLQSTSEARARYKFIEQQIGKRVDTFL 203  
 DB 121 IRENPLGLPALDSAITTLFFYNANSAASALMVLQSTSEARARYKFIEQQIGKRVDTFL 180  
 QY 204 PSLAIISLNSWSALSQKIQTASTNNGQFESPVLINAGNQRTVITNVDAGVVTSNIAL 263  
 DB 181 PSLAIISLNSWSALSQKIQTASTNNGQFETPVLINAGNQRTVITNVDAGVVTSNIAL 240  
 QY 264 LNRNNA 270  
 DB 241 PNRNNA 247  
 RESULT 7  
 ID Q00980 PRELIMINARY; PRT; 278 AA.  
 AC Q00980;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE B-luffin (EC 3.2.2.22) (rRNA N-glycosidase).  
 OS Luffa cylindrica (smooth loofah) (Sponge gourd).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids I; Cucurbitales; Cucurbitaceae; Luffa.  
 NCBI\_TaxID=3670;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92288316; PubMed=1600156;  
 RA Kataoka J., Habuka N., Miyano M., Masuta C., Koiwai A.;  
 RT "Nucleotide sequence of cDNA encoding alpha-luffin, a ribosome-  
 inactivating protein from Luffa cylindrica.";  
 RL Plant Mol. Biol. 18:1199-1202(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92353400; PubMed=1643290;  
 RA Kataoka J., Habuka N., Miyano M., Masuta C., Koiwai A.;  
 RT "Nucleotide sequence of cDNA encoding beta-luffin, another ribosome-  
 inactivating protein from Luffa cylindrica.";  
 RL Plant Mol. Biol. 19:887-889(1992).  
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 CC SPECIFIC ADENOSINE ON THE 28S RNA.  
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.

DR EMBL; X62372; CAA44230.1; -.  
 DR PIR; S23519; S23519.  
 DR HSP; P16094; 1AHC.  
 DR GO; GO:0016787; F:hydrolase activity; IEA.  
 DR GO; GO:0017149; F:protein biosynthesis inhibitor activity; IEA.  
 DR GO; GO:0030598; F:RNA N-glycosylase activity; IEA.  
 DR GO; GO:0015070; F:toxin activity; IEA.  
 DR GO; GO:0017148; P:negative regulation of protein biosynthesis; IEA.  
 DR InterPro; IPR001574; RIP.  
 DR Pfam; PF00161; RIP; 1.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
 KW Hydrolase; Toxin.  
 SQ SEQUENCE 278 AA; 30586 MW; 40C26576EB53F850 CRC64;  
 Query Match 54.1%; Score 775; DB 10; Length 278;  
 Best Local Similarity 58.6%; Pred. No. 3e-53;  
 Matches 163; Conservative 47; Mismatches 64; Indels 4; Gaps 3;  
 QY 1 MIRELVLSLILTLFTTPAVEG-DVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPL 59  
 DB 1 MNRFTFLSLILILIAFT--VEGANVSFSLSGADSKSYKFEITALKALPKSKVSNIP 58  
 QY 60 LRSLPSQRYALIHLTNYADETISVALDVTNVYMGYRAGDTSYFFNGASATEAAKYVF 119  
 DB 59 LLPSASGASRYLMQLSNYDAKAITMADVTNVYMGYRAGDTSYFFNGASATEAAKYVF 118  
 QY 120 KDMRKVTLPSYGNVERLQTAAGKIRENIPGLPALDSAITTLFFYNANSAASALMVLQ 179  
 DB 119 KGS-TIVTLPSYGNVERLQNAAGKREKIPGLFRAFDSAITSLFHYDSTAAGAPLVIIQ 177  
 QY 180 STBAARYKFIEQQIGKRVDTFLPSLAIISLNSWSALSQKIQTASTNNGQFESPVL 239  
 DB 178 TTAESAREKYEIQIIEIPKNEVPSPALSLNENWSALSQKIQTASTNNGQFESPVL 237  
 QY 240 NQGNORVITNVDAGVVTSNIALINRNNAAMDNDVP 277  
 DB 238 DMKGORVEIKDVKSNVNTNNIKLLNKONIAAFDDGIP 275  
 RESULT 8  
 ID Q9FUV7 PRELIMINARY; PRT; 286 AA.  
 AC Q9FUV7;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE MAP301 (EC 3.2.2.22) (rRNA N-glycosidase).  
 OS Momordica charantia (Bitter melon) (Balsam pear).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids I; Cucurbitales; Cucurbitaceae; Momordica.  
 NCBI\_TaxID=3673;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA QuanHong Y., Rihe P., Aisheng X.;  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 CC SPECIFIC ADENOSINE ON THE 28S RNA.  
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 DR EMBL; AF284811; AAG33028.1; -.  
 DR PIR; B61318; B61318.  
 DR HSP; P09989; 1MRJ.  
 DR GO; GO:0016787; F:hydrolase activity; IEA.  
 DR GO; GO:0017149; F:protein biosynthesis inhibitor activity; IEA.  
 DR GO; GO:0030598; F:RNA N-glycosylase activity; IEA.  
 DR GO; GO:0015070; F:toxin activity; IEA.  
 DR GO; GO:0017148; P:negative regulation of protein biosynthesis; IEA.  
 DR InterPro; IPR001574; RIP.  
 DR Pfam; PF00161; RIP; 1.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
 KW Hydrolase; Toxin.

SQ SEQUENCE 286 AA; 32031 MW; 5B2DF55A41D8F921 CRC64;

Query Match 52.9%; Score 758; DB 10; Length 286;  
 Best Local Similarity 59.2%; Pred. No. 6.9e-52;  
 Matches 157; Conservative 36; Mismatches 68; Indels 2; Gaps 2;

QY 1 MIREFLVSLLLTLPLTTTTPAVEGDVSRFLSGATSSYGVFISNRKALPNERKLYDIPL 60  
 Db 1 MVKCLLSFLIIAIFGVPTAKGVNFDLSTATAKTYTKFIEDFRATLPFSHKYDIPL 60

QY 61 RSLPSQSRVALHLNLYADETISVAIDVTNVIMGYRAGDTSYFENGASATAAKVYFK 120  
 Db 61 YSTISDSRRFILLNLTSYAYETISVAIDVTNVVWYAYTRDVSFFK-ESPPEAYNLFK 119

QY 121 DAMRKVTLPVSGNYERLQTAAGKIRENIPGLPALDSAITTLFFYNANSASALMVLIO 180  
 Db 120 -GTRKITLPTGYNENLQTAAGKIRENIDGLPALSSAITTLFFYNQAAPSALLVLIOT 178

QY 181 TSEARVYKFTQOIGKRVKDTFLPSLAISLNSWALSQKIOIASTNNQFESPVLIN 240  
 Db 179 TAEARFKYTERHVAKVATNFKPNLAISLENQWSALSQKIFLAQNGGKFRNPVDLIK 238

QY 241 AQNRVTITNDAGVVTNSIALLLN 265  
 Db 239 PTGERFQVTNVDSDVVKGNIKILLN 263

RESULT 9

Q41257 ID Q41257 PRELIMINARY; PRT; 286 AA.

AC Q41257; DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE MAP30 (BC 3.2.2.22) (rRNA N-glycosidase).  
 GN MAP30.

OS Momordica charantia (Bitter melon) (Balsam pear).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids I; Cucurbitales; Cucurbitaceae; Momordica.  
 OX NCBI\_TaxID=3673;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Leaf;  
 RX MEDLINE=95394347; PubMed=7665070;  
 RA Lee-Huang S., Huang P.L., Chen H.C., Huang P.L., Bourinbalar A.,  
 RT Huang H.I., Kung H.F.;  
 RT "Anti-HIV and anti-tumor activities of recombinant MAP30 from bitter  
 melon.";  
 RL Gene 161:151-156 (1995).

CC -|- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 CC SPECIFIC ADENOSINE ON THE 28S RRNA.

CC -|- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.

DR EMBL; S79450; AAB35194.2; -.  
 DR PIR; B61318; B61318.  
 DR PIR; J04235; J04235.  
 DR PDB; 1D8V; 10-MAY-00.  
 DR GO; GO:0016787; F:hydrolase activity; IEA.  
 DR GO; GO:0017149; F:protein biosynthesis inhibitor activity; IEA.  
 DR GO; GO:0030598; F:rRNA N-glycosylase activity; IEA.  
 DR GO; GO:0015070; F:toxin activity; IEA.  
 DR GO; GO:0017148; P:negative regulation of protein biosynthesis; IEA.  
 DR InterPro; IPR001574; RIP.  
 DR Pfam; PF00161; RIP; 1.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
 DR KW Hydrolase; Toxin.  
 DR FT NON TER 1 1  
 DR CHAIN 1 2 >264  
 FT NON TER 264 264  
 FT SEQUENCE 286 AA; 32018 MW; FA94E5A4A136F7C1 CRC64;

Query Match 52.6%; Score 753; DB 10; Length 286;  
 Best Local Similarity 58.9%; Pred. No. 1.7e-51;  
 Matches 156; Conservative 38; Mismatches 69; Indels 2; Gaps 2;

QY 1 MIREFLVSLLLTLPLTTTTPAVEGDVSRFLSGATSSYGVFISNRKALPNERKLYDIPL 60

Db 1 MVKCLLSFLIIAIFGVPTAKGVNFDLSTATAKTYTKFIEDFRATLPFSHKYDIPL 60

QY 61 RSLPSQSRVALHLNLYADETISVAIDVTNVIMGYRAGDTSYFENGASATAAKVYFK 120  
 Db 61 YSTISDSRRFILLNLTSYAYETISVAIDVTNVVWYAYTRDVSFFK-ESPPEAYNLFK 119

QY 121 DAMRKVTLPVSGNYERLQTAAGKIRENIPGLPALDSAITTLFFYNANSASALMVLIO 180  
 Db 120 -GTRKITLPTGYNENLQTAAGKIRENIDGLPALSSAITTLFFYNQAAPSALLVLIOT 178

QY 181 TSEARVYKFTQOIGKRVKDTFLPSLAISLNSWALSQKIOIASTNNQFESPVLIN 240  
 Db 179 TAEARFKYTERHVAKVATNFKPNLAISLENQWSALSQKIFLAQNGGKFRNPVDLIK 238

QY 241 AQNRVTITNDAGVVTNSIALLLN 265  
 Db 239 PTGERFQVTNVDSDVVKGNIKILLN 263

RESULT 10

Q9FSH2 ID Q9FSH2 PRELIMINARY; PRT; 264 AA.

AC Q9FSH2; DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Ribosome inactivating protein, RIP, type I (EC 3.2.2.22) (rRNA  
 DE N-glycosidase) (Fragment).  
 GN RIP.

OS Momordica charantia (Bitter melon) (Balsam pear).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids I; Cucurbitales; Cucurbitaceae; Momordica.  
 OX NCBI\_TaxID=3673;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Nguyen Huy H., Nohiem Ngoc M., Dao Huy P., Le Tran B., Nong Van H.;  
 RT "Expression of a RIP gene from Momordica charantia in E. coli.";  
 RL Submitted (SEP-2000) to the EMBL/Genbank/DBJ databases.

CC -|- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 CC SPECIFIC ADENOSINE ON THE 28S RRNA.

CC -|- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.

DR EMBL; AJ294541; CAC08217.1; -.  
 DR PIR; B61318; B61318.  
 DR HSP; P09899; IMR3.  
 DR GO; GO:0016787; F:hydrolase activity; IEA.  
 DR GO; GO:0017149; F:protein biosynthesis inhibitor activity; IEA.  
 DR GO; GO:0030598; F:rRNA N-glycosylase activity; IEA.  
 DR GO; GO:0015070; F:toxin activity; IEA.  
 DR GO; GO:0017148; P:negative regulation of protein biosynthesis; IEA.  
 DR InterPro; IPR001574; RIP.  
 DR Pfam; PF00161; RIP; 1.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
 DR KW Hydrolase; Toxin.  
 DR FT NON TER 1 1  
 DR CHAIN 1 2 >264  
 FT NON TER 264 264  
 FT SEQUENCE 264 AA; 29760 MW; AD9FCC032972718D CRC64;

Query Match 48.8%; Score 699; DB 10; Length 264;  
 Best Local Similarity 60.7%; Pred. No. 2.9e-47;  
 Matches 147; Conservative 31; Mismatches 62; Indels 2; Gaps 2;

QY 24 DVSRFLSGATSSYGVFISNRKALPNERKLYDIPLRSLPSQSRVALIHLNLYADETI 83  
 Db 2 DVNFDLSTATAKTYTKFIEDFRATLPFSHKYDIPLSYSTISDSRRFILLNLTSYAYETI 61

QY 84 SVAIDVTNVIMGYRAGDTSYFENGASATAAKVVFQDAMRKVTLTPVSGNYERLQTAAGK 143  
 Db 62 SVAIDVTNVVWYAYTRDVSFFK-ESPPEAYNLFK-GTRKITLPTGYNENLQTAAGK 119

144 IRENIPGLPALDSAITTLFFYNANSASALMVLIOSTSEARVKFIEQQIGKRVDTKTEL 203  
120 IRENIELGPAISAITTLFFYNQSPALLVLIQTAEARFKYIERHVKYVATNFK 179  
204 PSLAISLENSWSALSQIOIASTNNGQFSPVVLINAQORVITITNDAGVVTSTIAL 263  
180 PNLAIISLENSWSALSQIOIPLAQNGAKFRPVDLIKPTGERFQVTVNVDSDVKGNIKLL 239  
264 LN 265  
240 LN 241  
RESULT 11  
Q8SAD7 PRELIMINARY; PRT; 136 AA.  
AC Q8SAD7;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Type 1 ribosome-inactivating protein (EC 3.2.2.22) (rRNA  
N-glycosidase) (Fragment).  
OS Cucurbita moschata (Cushaw squash) (Winter crookneck squash).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC Cucurbitales; Cucurbitaceae; Cucurbita.  
OX NCBI\_TaxID=3662;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Lin Y., Wu Z.J., Lin Q.Y., Xie L.H.;  
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
CC -!- CATALYTIC ACTIVITY: ENDOLYOLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
SPECIFIC ADENOSINE ON THE 28S RRNA.  
CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
DR EMBL; AF462349; AAL67855.2; -  
DR GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.  
DR GO; GO:0017149; F:protein biosynthesis inhibitor activity; IEA.  
DR GO; GO:0030598; F:RNA N-glycosylase activity; IEA.  
DR GO; GO:0030598; F:RNA N-glycosylase activity; IEA.  
DR GO; GO:0015070; F:toxin activity; IEA.  
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.  
DR GO; GO:0017148; P:negative regulation of protein biosynthesis; IEA.  
DR InterPro; IPR001574; RIP.  
DR Pfam; PF00161; RIP; 1.  
DR PRINTS; PR00396; SHIGARICIN.  
DR Glycosidase; Hydrolase; Toxin.  
FT NON\_TER 1  
FT NON\_TER 136  
SQ SEQUENCE 136 AA; 15195 MW; F728C9E80B176D61 CRC64;  
Query Match 32.8%; Score 469.5; DB 10; Length 136;  
Best Local Similarity 67.9%; Pred. No. 1.6e-29;  
Matches 93; Conservative 21; Mismatches 22; Indels 1; Gaps 1;  
QY 89 VTNYVINGRAGTSTYFFNGASATEAAKYVFKDAMRKVTLPYSGNYERLQTAAGKIRENI 148  
Db 1 VTNYVINGVLYNSTSTFFNEDSQAQSKYVFKDSTR-ITLPYSGNYEKLQAAGKKREKI 59  
QY 149 PLGLPALDSAITTLFFYNANSASALMVLIOSTSEARVKFIEQQIGKRVDTKTELPSLAI 208  
Db 60 PLGLPALDNAITTLFHYDSTAAAAFTWIIQCTAEASRYRYIEAQMIKRISKDDVPSLAI 119  
QY 209 ISLENSWSALSQIOIA 225  
Db 120 ISLENSWSALSQIOIFA 136  
RESULT 12  
Q8GT32 PRELIMINARY; PRT; 563 AA.  
AC Q8GT32;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DE Ribosome-inactivating protein (EC 3.2.2.22) (rRNA  
N-glycosidase).  
OS Sambucus nigra (European elder).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
OC Campanulids; Dipsacales; Adoxaceae; Sambucus.  
OX NCBI\_TaxID=4202;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA TISSUE=Leaf;  
RC "Characterization and molecular cloning of Nigrin 1, a type two  
ribosome-inactivating protein from leaves of elder (Sambucus nigra).";  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF249280; AAN86130.1; -  
DR GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.  
DR GO; GO:0017149; F:protein biosynthesis inhibitor activity; IEA.  
DR GO; GO:0030598; F:RNA N-glycosylase activity; IEA.  
DR GO; GO:0015070; F:toxin activity; IEA.  
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.  
DR GO; GO:0017148; P:negative regulation of protein biosynthesis; IEA.  
DR InterPro; IPR008997; RicinB-like.  
DR InterPro; IPR000772; RicinB-lectin.  
DR InterPro; IPR001574; RIP.  
DR Pfam; PF00652; RicinB-lectin; 6.  
DR Pfam; PF00161; RIP; 1.  
DR PRINTS; PR00396; SHIGARICIN.  
DR SMART; SM00458; RICIN; 2.  
DR PROSITE; PS00231; RICIN\_B-LECTIN; 2.  
DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
KW Hydrolase; Glycosidase.  
SQ SEQUENCE 563 AA; 62173 MW; 0EB236421FC5E04F CRC64;  
Query Match 30.6%; Score 437.5; DB 10; Length 563;  
Best Local Similarity 36.2%; Pred. No. 4.2e-26;  
Matches 102; Conservative 66; Mismatches 101; Indels 13; Gaps 6;  
QY 6 VLSLLILTLFTTPAVEG---DVSFRLSGATSSYGVFISNLKAL-PNERKLYDIPLL 60  
Db 8 MLFYFIVLVAICSVGIQIDYPSVFNLDGAKSATVRDPLSLNRKTVATGTVEVGLPVL 67  
QY 61 R--SSLPQSQRVALIHLTYADETISVAIDVTNVIIMGVYRAGDTSYFFNGASATEAAKYV 118  
Db 68 RRESEVQVKSREVLPLTNNGTNTVLAVDVNLVYVAFSGNANSYFFK--DATEVQKSN 125  
QY 119 FKDAMRKVTLPYSGNYERLQTAAGKIRENIPLGLPALDSAITTLFFYNANSASALMVL 178  
Db 126 LFVGTQNTLSFGVNDNLETAANTFERESIEGLGPSPLDGAITSL--YHGDSVARSLVVI 183  
QY 179 QSTSEAAKYFIEQQIGKRVDK--TFPLSLAISLENSWSALSQIOIASTNNGQFSPV 236  
Db 184 QMVSEAAFRYIEQEVRRSLQQATSTFPNASMLSMENNWSMSLEIQAGNNVSPFSGTV 243  
QY 237 VLINAQORVITITNDAGVVTSTIALLNENMAAMDDDDVPM 278  
Db 244 QLINYDTHRLVDNPFELYKITGIAILLFRCSPPSNDNAIRM 285  
RESULT 13  
Q945S2 PRELIMINARY; PRT; 563 AA.  
AC Q945S2;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Ribosome-inactivating protein (EC 3.2.2.22) (rRNA  
N-glycosidase).  
OS Sambucus nigra (European elder).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
OC Campanulids; Dipsacales; Adoxaceae; Sambucus.

OX NCBI\_TaxID=4202;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Leaf;  
 RA Van Damme E.J.M.;  
 RT "Characterization and cloning of lectins and ribosome-inactivating  
 proteins from *Sambucus nigra* leaves.";  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 CC SPECIFIC ADENOSINE ON THE 28S RRNA.  
 DR EMBL; U76524; AAC15886.1; -  
 DR HSP; P02879; 2AAL  
 DR GO; GO:0016787; F:hydrolase activity; IEA.  
 DR GO; GO:0017449; F:protein biosynthesis inhibitor activity; IEA.  
 DR GO; GO:0030598; F:RNA N-glycosylase activity; IEA.  
 DR GO; GO:0015070; F:toxin activity; IEA.  
 DR GO; GO:0017148; P:negative regulation of protein biosynthesis; IEA.  
 DR InterPro; IPR008997; RicinB like.  
 DR InterPro; IPR000772; RicinB\_lectin.  
 DR Pfam; PF00652; RicinB\_lectin; 6.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR SMART; SM00458; RICIN; 2.  
 DR PROSITE; PS0231; RICIN\_B\_LECTIN; 2.  
 DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
 KW Hydrolase; Toxin.  
 SQ SEQUENCE 563 AA; 62242 MW; 07F7CBEDCF33BF10 CRC64;  
 Query Match 30.3%; Score 433.5; DB 10; Length 563;  
 Best Local Similarity 35.9%; Pred. No. 8.8e-26;  
 Matches 103; Conservative 66; Mismatches 105; Indels 13; Gaps 6;  
 QY 1 MIRFLVLSLLILTLFTTTPAVEG----DVSFRISGATSSSYGVFISNLRKAL-PNERKLY 55  
 Db 3 VIAAAMLYFYVVLVLAICSVGIQIDYPSVSNFLDGAKSATYRDFLNLRKTIVATGYEVN 62  
 QY 56 DIPLLR--SSLPQSRYALIHNTYADETISVAIDVTNVTYIMGYRAGDTSYFNGASATE 113  
 Db 63 GLPVLRESEVQVKRFVLPLTNNGTNTVLAVDVTNLYVVAFSNANSYFFK--DATE 120  
 QY 114 AAKYVFKDAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFFYNANSAASA 173  
 Db 121 VQKSNLFVCTKNTLSFTGNYDNLETAAGTRESIELGSPDLGATSL--YHGDSVARS 178  
 QY 174 LMVLIQSTSEAAARYKFIQOIGKRVDK--TFLPSLAIISLNSWSALSQIOIASTNNQ 231  
 Db 179 LLVVIQMVSEAAARFYIEQEVRSRLOQATSFPTNALMLSWENNWSMSLEIQAGNNVSP 238  
 QY 232 FESPVLINAQORVTITNDAGVVTSTNIALLRNNMAAMDDVPM 278  
 Db 239 PFGTVOLLNVDYHTRLVDFNFEELYKITGIAILLFRCSPPSNDNAIRM 285  
 RESULT 14  
 ID O04367 PRELIMINARY; PRT; 563 AA.  
 AC O04367;  
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)  
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Ribosome inactivating protein precursor (BC 3.2.2.22) (rRNA  
 DE N-glycosidase).  
 OS *Sambucus nigra* (European elder).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; ascerids;  
 OC campanulids; Dipsacales; Adoxaceae; Sambucus.  
 OX NCBI\_TaxID=4202;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98112023; PubMed=9450339;  
 RA Van Damme E.J., Roy S., Barre A., Ronge P., Van Leuven F.,  
 RA Peumans W.J.;

"The major elderberry (*Sambucus nigra*) fruit protein is a lectin  
 derived from a truncated type 2 ribosome-inactivating protein.";  
 RL Plant J. 12:1251-1260(1997).  
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 CC SPECIFIC ADENOSINE ON THE 28S RRNA.  
 DR EMBL; U76524; AAC15886.1; -  
 DR HSP; P02879; 2AAL  
 DR GO; GO:0016787; F:hydrolase activity; IEA.  
 DR GO; GO:0017449; F:protein biosynthesis inhibitor activity; IEA.  
 DR GO; GO:0030598; F:RNA N-glycosylase activity; IEA.  
 DR GO; GO:0015070; F:toxin activity; IEA.  
 DR GO; GO:0017148; P:negative regulation of protein biosynthesis; IEA.  
 DR InterPro; IPR008997; RicinB like.  
 DR InterPro; IPR000772; RicinB\_lectin.  
 DR Pfam; PF00652; RicinB\_lectin; 6.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR SMART; SM00458; RICIN; 2.  
 DR PROSITE; PS0231; RICIN\_B\_LECTIN; 2.  
 DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
 KW Hydrolase; Signal; Toxin.  
 FT SIGNAL 1 25 POTENTIAL.  
 FT CHAIN 26 297 RIBOSOME INACTIVATING PROTEIN, A CHAIN.  
 FT CHAIN 298 563 RIBOSOME INACTIVATING PROTEIN, B CHAIN.  
 SQ SEQUENCE 563 AA; 62336 MW; 3ED2BC08E796205 CRC64;  
 Query Match 30.1%; Score 430.5; DB 10; Length 563;  
 Best Local Similarity 37.1%; Pred. No. 1.5e-25;  
 Matches 102; Conservative 62; Mismatches 98; Indels 13; Gaps 6;  
 QY 1 MIRFLVLSLLILTLFTTTPAVEG----DVSFRISGATSSSYGVFISNLRKAL-PNERKLY 55  
 Db 3 VIAAAMLYFYVVLVLAICSVGIQIDYPSVSNFLDGAKSATYRDFLNLRKTIVATGYEVN 62  
 QY 56 DIPLLR--SSLPQSRYALIHNTYADETISVAIDVTNVTYIMGYRAGDTSYFNGASATE 113  
 Db 63 GLPVLRESEVQVKRFVLPLTNNGTNTVLAVDVTNLYVVAFSNANSYFFK--DATE 120  
 QY 114 AAKYVFKDAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFFYNANSAASA 173  
 Db 121 LQKSNLFVCTKNTLSFTGNYDNLETAAGTRESIELGSPDLGATSLY--DESVARS 178  
 QY 174 LMVLIQSTSEAAARYKFIQOIGKRVDK--TFLPSLAIISLNSWSALSQIOIASTNNQ 231  
 Db 179 LLVVIQMVSEAAARFYIEQEVRSRLOQATSFPTNALMLSMENNWSMSLEVQOQGNVSP 238  
 QY 232 FESPVLINAQORVTITNDAGVVTSTNIALLR 266  
 Db 239 FGTGTVOLLNVDYHTRLVDFNFEELYKITGIAILLFR 273  
 RESULT 15  
 ID Q41174 PRELIMINARY; PRT; 541 AA.  
 AC Q41174;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Proridin A chain (EC 3.2.2.22) (rRNA N-glycosidase)  
 DE (Fragment).  
 OS *Ricinus communis* (Castor bean).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids I; Malpighiales; Euphorbiaceae; Acalyphoideae; Acalyphaceae;  
 OC Ricinus.  
 OX NCBI\_TaxID=3988;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92338377; PubMed=1633311;  
 RA Roberts L.M., Tregear J.W., Lord J.M.;  
 RA "Molecular cloning of ricin.";



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 14, 2004, 18:55:11 ; Search time 3322 Seconds  
(without alignments)  
3770.661 Million cell updates/sec

Title: US-09-905-247-1  
Perfect score: 1432  
Sequence: 1 MRFVLVLSLLILTLFTTPA.....AAMDVDPMTQSGFGSGVAL 289

Scoring table: BLOSUM62 Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool\_p/US09905247/runat\_10092004\_103933\_4713/app\_query.fasta\_1.455  
-DB=GenEmbl -QFMT=fastp -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09905247 @CGN 1 1 5600 @runat\_10092004\_103933\_4713 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl :  
1: gb\_ba :  
2: gb\_htg :  
3: gb\_in :  
4: gb\_om :  
5: gb\_ov :  
6: gb\_pat :  
7: gb\_ph :  
8: gb\_pl :  
9: gb\_pr :  
10: gb\_ro :  
11: gb\_sts :  
12: gb\_sy :  
13: gb\_un :  
14: gb\_vi :  
15: em\_ba :  
16: em\_fun :  
17: em\_hum :  
18: em\_in :  
19: em\_mu :  
20: em\_om :  
21: em\_or :  
22: em\_ov :  
23: em\_pat :  
24: em\_ph :  
25: em\_pl :  
26: em\_ro :  
27: em\_sts :  
28: em\_un :

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1422	99.3	1010	8	TRHTCSA	M34858 T.kirillowii
2	1419	99.1	1809	8	AY082349	AY082349 Trichosan
3	1418	99.0	1799	8	AF367252	AF367252 Trichosan
4	1413	98.7	886	6	AR032228	AR032228 Sequence 3
5	1413	98.7	886	6	I32881	I32881 Sequence 3
6	1413	98.7	886	6	AR228702	AR228702 Sequence
7	1413	98.7	886	6	BD011600	BD011600 Recombina
8	1413	98.7	886	6	BD165859	BD165859 Transgeni
9	1413	98.7	1227	8	TRHTCS	J05434 T.kirillowii
10	1408	98.3	1134	8	S70176	S70176 trichosanth
11	1377	96.2	1194	8	AB000666	AB000666 Trichosan
12	1262	88.1	810	8	TKU25675	U25675 Trichosan
13	1253	87.5	1433	8	AY082348	AY082348 Trichosan
14	1180	82.4	741	8	AB039324	AB039324 Trichosan
15	1164.5	81.3	1084	8	BYNBD1	BYNBD1 Trichosan
16	1164.5	81.3	1094	6	I24020	I24020 Bryonia dio
17	1074	75.0	1499	6	AR117137	AR117137 Sequence 1
18	1064	74.3	758	6	AR117133	AR117137 Sequence 1
19	859.5	60.0	1044	8	MCAMC	MCAMC
20	802.5	56.0	1026	6	E03988	E03988 M.charantia
21	802.5	56.0	1026	6	E03988	E03988 DNA encodin
22	775	54.1	914	8	LCALUF	LCALUF
23	775	54.1	914	8	E03989	E03989 DNA encodin
24	766.5	53.5	1012	6	LCBLUF	LCBLUF
25	758	52.9	861	8	AF284811	AF284811 Momordica
26	756	52.8	914	8	MBMOMORII	MBMOMORII
27	753	52.6	919	8	S79450	S79450 Map30=anti-
28	736	51.4	1042	8	AB045560	AB045560 Cucumis f
29	728	50.8	913	6	AR003723	AR003723 Sequence
30	728	50.8	913	6	AR010059	AR010059 Sequence
31	728	50.8	913	6	AR055265	AR055265 Sequence
32	728	50.8	913	6	AR141194	AR141194 Sequence
33	728	50.8	913	6	AR141431	AR141431 Sequence
34	728	50.8	913	6	I11903	I11903 Sequence 13
35	728	50.8	913	6	I40476	I40476 Sequence 13
36	728	50.8	913	6	AR367960	AR367960 Sequence
37	728	50.8	913	6	AR430941	AR430941 Sequence
38	699	48.8	795	8	MCH294541	MCH294541 Momordica
39	685.5	47.9	1039	8	AF055086	AF055086 Trichosan
40	537.5	37.5	962	6	I34238	I34238 Sequence 14
41	469.5	32.8	408	8	AF462349	AF462349 Cucurbita
42	437.5	30.6	1906	8	AF249280	AF249280 Sambucus
43	433.5	30.3	1692	8	AF409135	AF409135 Sambucus
44	432.5	30.2	1879	8	SNU41299	SNU41299 Sambucus ni
45	430.5	30.1	2079	8	SNU76524	U76524 Sambucus ni

ALIGNMENTS

TRHTCSA 1010 bp mRNA linear PLN 27-APR-1993  
 LOCUS T.kirilowii trichosanthin (TCS) mRNA, complete cds.  
 DEFINITION M34858  
 ACCESSION M34858  
 VERSION 1  
 KEYWORDS ribosome-inactivating protein; trichosanthin.  
 SOURCE Trichosanthin kirilowii (Mongolian snake-gourd)  
 ORGANISM Trichosanthin kirilowii  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthin.  
 1 (bases 1 to 1010)  
 REFERENCE 1  
 AUTHORS Shaw, P.-C., Yung, M.H., Zhu, R.H., Ho, W.K., Ng, T.B. and Yeung, H.W.  
 TITLE Cloning of trichosanthin cDNA and its expression in *Escherichia coli*  
 JOURNAL Gene 97 (2), 267-272 (1991)  
 MEDLINE 91153657  
 PUBMED 1999291  
 COMMENT Original source text: T.kirilowii maximowicz, cDNA to mRNA.  
 Draft entry and computer-readable sequence for [Unpublished (1990)]  
 kindly submitted  
 by P.-C.Shaw, 06-JUN-1990.  
 Author address: P.-C.Shaw  
 Department of Biochemistry  
 Chinese University of Hong Kong  
 Shatin, NT, HONG KONG.  
 FEATURES  
 source Location/Qualifiers  
 1..1010  
 /organism="Trichosanthin kirilowii"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3677"  
 10..879  
 /note="Trichosanthin precursor"  
 /codon\_start=1  
 /protein\_id="AAA34207.1"  
 /db\_xref="GI:170537"  
 /translation="MIRFLVSLTLTLFTTPAVEGDVSRSLGATSSSYGVFISNL  
 RKALNERKLDIPIRLSRSLPGSRVLIHLTNVADETISVAIDVTNVIMYGRAGDT  
 SYFFNEASATBAKVFDDMKKVTLPYSGNVERLOTRAGKIRENIPGLPALDSAIT  
 TLFYNNASASALMVLIOSTSEAAKYFIEQIGKRVDKTLPPLSLAISLNSWAL  
 PKQIQIASTNNQGFSPVLLINAQNRVTITNVDAGVVTNSIALLLNNRNNMAAMDDVD  
 PMQTQFGGGSYAI"  
 10..78  
 /note="trichosanthin signal peptide"  
 79..819  
 /product="trichosanthin"  
 196  
 /note="t in wild type; a in allele"  
 197  
 /note="c in wild type; g in allele"  
 468  
 /note="t in wild type; a in allele"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 5,15e-120 Length: 1010  
 Score: 1422.00 Matches: 287  
 Percent Similarity: 99.65% Conservative: 1  
 Best Local Similarity: 99.31% Mismatches: 1  
 Query Match: 99.30% Indels: 0  
 DB: 8 Gaps: 0  
 US-09-905-247-1 (1-289) x TRHTCSA (1-1010)  
 QY 1 MetileArgPheLeuValLeuSerLeuLeuLeuThrLeuThrLeuThrProAla 20  
 Db 10 ATGATCAGATTCTAGTCTCTCTTCTGTAATCTCACCTCTCTCTCAACACTCTGCT 69  
 QY 21 ValGluGlyAspValSerPheArgLeuSerGlyValAlaThrSerSerTyrGlyValPhe 40  
 Db 70 GTGAGGGCGATGTAGCTTCGGTTTATCAGGTGCAACAGAGTCTCTATGAGTTTC 129  
 QY 41 IleSerAsnLeuArgLysAlaLeuProAsnGluArgLysLeuTyrAspIleProLeuLeu 60

130 ATTTCAAATCTGAGAAAGCTCTTCCAAATGAAAGGAACTGTACGATATCCCTCTCTTTA 189  
 QY 61 ArgSerSerLeuProGlySerGlnArgTyrAlaLeuLeuHisLeuThrAsnTyrAlaAsp 80  
 Db 190 CGTTCTCTCTTCCAGGTTCTCAACGCTAGCATTTGATTCATCTCACAAATATTAGCCGAT 249  
 QY 81 GluThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGly 100  
 Db 250 GAAACCATTTTCAGTGGCCATAGACGTAAACGACGCTATATATATGGGATATCGCGTGGC 309  
 QY 101 AspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPheLys 120  
 Db 310 GATACATCTATTTTTCACAGAGGCTTCTGCAACAGAAAGCTGCAAAATATGTTATTCAA 369  
 QY 121 AspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAla 140  
 Db 370 GACGCTATGCGAAAGTTACGCTTCATATCTGCAATATTACGAAGGCTTCAAAATGCT 429  
 QY 141 AlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThr 160  
 Db 430 GCAGGCAAAATAGGGAATAATATTCGCTTGGACTCCCTGCTTTGGACAGTGCATATACC 489  
 QY 161 ThrLeuPheTyrTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSer 180  
 Db 490 ACTTTGTTTACTACAACGCCAATTCGCTGCTGGCATTATGCTACTCATTCAGTCG 549  
 QY 181 ThrSerGluAlaAlaArgTyrLysPheIleGluGlnGlnIleGlyLysArgValAspLys 200  
 Db 550 ACGTCTGAGGCTGCGAGGTATTAATTTATTGAGCAACAATTTGGGAACGCTTGCACAAA 609  
 QY 201 ThrPheLeuProSerLeuAlaIleSerLeuGluAsnSerTrpSerAlaLeuSerLys 220  
 Db 610 ACCTTCTTACCAAGTTTAGCAATTAAGTTTGGAAATAGTTGCTGCTCTCTCTCCAAG 669  
 QY 221 GlnIleGlnIleAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuIleAsn 240  
 Db 670 CAAATTCAGATAGCAGTACTAATAATGACAGATTGAAAGTCTCTGTTGCTTTAAT 729  
 QY 241 AlaGlnAsnGlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnIle 260  
 Db 730 GCTCAAAACCAACAGAGTACGATACCAATGTTGATGCTGGAGTTGTAACCTTCAACATC 789  
 QY 261 AlaLeuLeuLeuAsnArgAsnAsnMetAlaAlaMetAspAspValProMetThrGln 280  
 Db 790 GCGTTGCTGCTGAATAGAAACAATATGCGACGATGATGACGATGTTCTTATGACACAG 849  
 QY 281 SerPheGlyCysGlySerTyrAlaLeu 289  
 Db 850 AGCTTTGGATGTGGAAGTTATGCTATT 876  
 RESULT 2  
 LOCUS AY082349 1809 bp DNA linear PLN 24-MAR-2003  
 DEFINITION Trichosanthin kirilowii trichosanthin precursor (TCS) gene,  
 complete cds.  
 ACCESSION AY082349  
 VERSION AY082349.2 GI:29165640  
 KEYWORDS Trichosanthin kirilowii (Mongolian snake-gourd)  
 SOURCE Trichosanthin kirilowii  
 ORGANISM Trichosanthin kirilowii  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthin.  
 1 (bases 1 to 1809)  
 REFERENCE 1  
 AUTHORS Yuan, H., Wang, L., Wang, Y., Liu, T., An, C. and Chen, Z.  
 TITLE Trichosanthin kirilowii trichosanthin precursor (TCS) gene  
 JOURNAL Unpublished  
 REFERENCE 2  
 AUTHORS Yuan, H., Wang, L., Wang, Y., Liu, T., An, C. and Chen, Z.  
 TITLE Direct Submission  
 JOURNAL Submitted (05-MAR-2002) National Laboratory of Protein Engineering  
 and Plant Genetic Engineering, College of Life Sciences, Peking



Db	1420	ACTTGTTGTTTACTACACGCAATTCCTGCTGCGCTATATGGTACTCAATTCAGTCG	1479
Qy	181	ThrSerGluAlaAlaArgTyrLysPheIleGluGlnIleGlyLysArgValAspLys	200
Db	1480	ACGTCTGAGCTCGGAGGTATAAATTATTGACACAAATTCGGAGCGGTGTGACAAA	1539
Qy	201	ThrPheLeuProSerLeuAlaIleIleSerLeuGluAsnSerTrpSerAlaLeuSerLys	220
Db	1540	ACCTTCTACCAAGTTAGCAATTATAAGTTTGAAAAATAGTTGGTCTGCTCTCTCCAAG	1599
Qy	221	GlnIleGlnIleAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuIleAsn	240
Db	1600	CAAAATTCAGATACGAGTACTAATAATGACACAGTTTGAAAGTCCTGTTGGCTTATAAT	1659
Qy	241	AlaGlnAsnGlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnIle	260
Db	1660	GCTCAAAACCAACGAGTCAGATAACCAATGTTGATGCTGGAGTTGTAACCTCCAAATC	1719
Qy	261	AlaLeuLeuLeuAsnArgAsnAsnMetAlaAlaMetAspAspValProMetThrGln	280
Db	1720	GCCTTGCTGCTGAATAGAACCAATATGGCAGCCATGATGACGATGTTCTATGACACAG	1779
Qy	281	SerPheGlyCysGlySerTyrAlaLeu	289
Db	1780	AGCTTTGGATGTGGAAGTTATGCTATT	1806
RESULT 3			
AF367252			
LOCUS			
DEFINITION	AF367252	1799 bp DNA linear PLN 14-MAY-2002	
ACCESSION	AF367252	Trichosanthes kirilowii trichosanthin precursor (TCS) gene,	
VERSION	AF367252	complete cds.	
KEYWORDS	AF367252.1	GI:14030568	
SOURCE			
ORGANISM		Trichosanthes kirilowii (Mongolian snake-gourd)	
REFERENCE		Trichosanthes kirilowii	
AUTHORS		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
TITLE		Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;	
JOURNAL		rosids; eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.	
FEATURES		1 (bases 1 to 1799)	
source		Yuan,H., Wang,L., Wang,Y., An,C. and Chen,Z.	
		Direct Submission	
		Submitted (31-MAR-2001) National Laboratory of Protein Engineering	
		and Plant Genetic Engineering, College of Life Sciences, Peking	
		University, Beijing 100871, China	
		Location/Qualifiers	
		1..1799	
		/organism="Trichosanthes kirilowii"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:3677"	
		/clone="type2"	
gene		1..>1799	
		/gene="TCS"	
promoter		1..904	
		/gene="TCS"	
misc_feature		903	
		/gene="TCS"	
		/note="alternative weaker transcriptional start site"	
mRNA		905..>1799	
		/gene="TCS"	
5'UTR		905..929	
		/product="trichosanthin precursor"	
		/gene="TCS"	
misc_feature		906	
		/gene="TCS"	
		/note="alternative weaker transcription start site"	
CDS		930..1799	
		/gene="TCS"	
		/function="inactivates eukaryotic 60S ribosomal subunits"	
		/note="preproprotein"	
		/codon_start=1	
		/product="trichosanthin precursor"	

AR032228	AR032228	Sequence 3 from patent US 5866785.	886 bp	DNA	linear	PAT 29-SEP-1999
LOCUS	AR032228					
DEFINITION	Sequence 3 from patent US 5866785.					
ACCESSION	AR032228					
VERSION	AR032228.1	GI:5946517				
KEYWORDS	.					
SOURCE	Unknown.					
ORGANISM	Unknown.					
REFERENCE	Unclassified.					
AUTHORS	1 (bases 1 to 886)					
TITLE	Donson, J., Dawson, W.O., Grantham, G.L., Turpen, T.H., Turpen, A., Myers, G., Garger, S.J., and Grill, L.K.					
JOURNAL	Recombinant plant viral nucleic acids					
FEATURES	Patent: US 5866785-A 3 02-FEB-1999;					
source	Location/Qualifiers					
	1..886					
	/organism="unknown"					
	/mol_type="unassigned DNA"					
ORIGIN						
Alignment Scores:						
Pred. No.:	2.85e-119	Length:	886			
Score:	1413.00	Matches:	285			
Percent Similarity:	99.31%	Conservative:	2			
Best Local Similarity:	98.82%	Mismatches:	0			
Query Match:	98.67%	Indels:	0			
DB:	6	Gaps:	0			
US-09-905-247-1 (1-289) x AR032228 (1-886)						
Qy	1 MettleArgPheLeuValLeuSerLeuLeuLeuThrLeuPheLeuThrThrProAla	20				
Db	8 ATGATCAGATTCTTAGTCCTCTCTTTGGTAAATTCACCTCTTCTCTAACAACTCTGCT	67				
Qy	21 ValGluGlyAspValSerPheArgLeuSerGlyAlaThrSerSerSerTyrGlyValPhe	40				
Db	68 GTGGAGGGCGATGTTAGCTTCGGTTTATCAGGTGCAACAGCAGTTCCTATGGAGTTTTC	127				
Qy	41 IleSerAsnLeuArgLysAlaLeuProAsnGluArgLysLeuTyrAspIleProLeuLeu	60				
Db	128 ATTTCAAATCTGAGAAAAGCTCTTCCAATGAAGGAAACCTGACGATATCCTCTGTTA	187				
Qy	61 ArgSerSerLeuProGlySerGlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAsp	80				
Db	188 CGTTCTCTCTTCCAGGTTCTCAAGCTACGATTTGATCCATCTCACAAATACGCCGAT	247				
Qy	81 GluThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGly	100				
Db	248 GAAACCAATTCAGTGGCCATAGACGTACGAAACGCTATATATATGGGATATCGCGCTGC	307				
Qy	101 AspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPheLys	120				
Db	308 GATACATCTATTTTTTCAACGAGGCTTCTGCAACAGAAAGCTTGCAAAATATGTATTCAA	367				
Qy	121 AspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAla	140				
Db	368 GACGCTATCGGAAAAGTTACGCTTCCATATTTCTGGCAATACGAAAAGGCTTCAAACCTGCT	426				

```
QY 141 AlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThr 160
Db 428 GCGGGCAAAATAAGGAAATAATTCGCTGGACTCCCGCTTGGACATGGCATTACC 487
QY 161 ThrLeuPheTyrTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSer 180
Db 488 ACTTTGTTTACTACAAACGCAATTCCTGCTGCGCTCGGCATCTATGGTACTCAITTCAGTCG 547
QY 181 ThrSerGluAlaAlaArgTyrLysPheIleGluGlnGlnIleGlyLysArgValAspLys 200
Db 548 ACGTCTGAGGCTCGAGAGTATAAATTTATTTAGCAACAATAATCGGAGCGCGTTGACAAA 607
QY 201 ThrPheLeuProSerLeuAlaIleIleSerLeuGluAsnSerTyrSerAlaLeuSerLys 220
Db 608 ACCCTTCTACCAAGTTAGCAATATAAGTTTGGAAATAGTTGGTCTCTCTCCAG 667
QY 221 GlnIleGlnIleAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuIleAsn 240
Db 668 CAAATTCAGATACGAGTACTAATAATATGGACAGTTTGAACCTCCTGTGTGCTTATAAAT 727
QY 241 AlaGlnAsnGlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnIle 260
Db 728 GCTCAAAACCAACGAGTCATGATAACCAATGTTGATGCTGGAGTTGTAACTTCAACATC 787
QY 261 AlaLeuLeuLeuAsnArgAsnAsnMetAlaAlaMetAspAspValProMetThrGln 280
Db 788 GCGTTGCTGCTGAATCGAAACAATATGGCAGCCATGGATGACGATGTTCTATGACACAG 847
QY 281 SerPheGlyCysGlySerTyrAlaLeu 289
Db 848 AGCTTTGGATGTGGAAGTTATGCTATT 874

RESULT 5
LOCUS I32881 886 bp DNA linear PAT 06-FEB-1997
DEFINITION Sequence 3 from patent US 5589367.
ACCESSION I32881
VERSION I32881.1 GI:1823672
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Unclassified.
1 (bases 1 to 886)
Donson,J., Dawson,W.O., Grantham,G.L., Turpen,T.H., Turpen,A.M.,
Garger,S.J. and Grill,L.K.
TITLE Recombinant plant viral nucleic acids
JOURNAL Patent: US 5589367-A 3 31-DEC-1996;
FEATURES
source
1..886
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Alignment Scores:
Pred. No.: 2.85e-119 Length: 886
Score: 1413.00 Matches: 285
Percent Similarity: 99.31% Conservative: 2
Best Local Similarity: 98.62% Mismatches: 2
Query Match: 98.67% Indels: 0
DB: 6 Gaps: 0

US-09-905-247-1 (1-289) x I32881 (1-886)

QY 1 MetIleArgPheLeuValLeuSerLeuLeuIleLeuThrLeuPheLeuThrThrProAla 20
Db 8 ATGATCAGATCTTAGTCTCTCTTGTGTAATCTCAACCTCTTCTTAACTCTCTGCT 67
QY 21 ValGluGlyAspValSerPheArgLeuSerGlyAlaThrSerSerSerTyrGlyValPhe 40
Db 68 GTGAGGGCGGATGTTAGTCTCGTATTACAGTGCACAACAGCTTCCTATGGAGTTTTC 127
QY 41 IleSerAsnLeuArgLysAlaLeuProAsnGluArgLysLeuTyrAspIleProLeuLeu 60
```

```
Db 128 ATTTCAAATCTGAAAAAGCTCTTCCAAATGAAGGAAACTGTACGATATCCCTCTCTTA 187
QY 61 ArgSerSerLeuProGlySerGlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAsp 80
Db 188 CGTTCTCTCTTCCAGGTTCTCAACGCTACGCAATGATCCATCTCACAANTTACGCCGAT 247
QY 81 GluThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGly 100
Db 248 GAAACCATTTTCAGTGCCATAGACGTACACAACTGCTATATTATGGGATATCGCGTGGC 307
QY 101 AspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPheLys 120
Db 308 GATACATCTCTATTTTTCACGAGGCTTCTGCACAGAAAGCTCAAAATATGATTACAA 367
QY 121 AspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAla 140
Db 368 GACGCTATCGAAAAGTTACGCTTCCATATTCCTGGCAATTCGAAAGGCTTCAACTGCT 427
QY 141 AlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThr 160
Db 428 GCGGGCAAAATAAGGAAATAATTCGCTTGGACTCCCGCTTGGACAGTGCATTACC 487
QY 161 ThrLeuPheTyrTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSer 180
Db 488 ACTTTGTTTACTACAAACCCCAATTCCTGCTGCGCACTTATGTTACTCATCTCAGTCG 547
QY 181 ThrSerGluAlaAlaArgTyrLysPheIleGluGlnGlnIleGlyLysArgValAspLys 200
Db 548 ACGTCTGAGGCTCGCAGGTATAAATTTATTTAGCAACAATAATGGGAGCGCGTTGACAA 607
QY 201 ThrPheLeuProSerLeuAlaIleIleSerLeuGluAsnSerTyrSerAlaLeuSerLys 220
Db 608 ACCCTTCTACCAAGTTTAGCAATATAAGTTTGGAAATAGTTGGTCTCTCTCTCCAAG 667
QY 221 GlnIleGlnIleAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuIleAsn 240
Db 668 CAAATTCAGATACGAGTACTAATAATATGGACAGTTTGAACCTCCTGTGTGCTTATAAAT 727
QY 241 AlaGlnAsnGlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnIle 260
Db 728 GCTCAAAACCAACGAGTCATGATAACCAATGTTGATGCTGGAGTTGTAACTTCAACATC 787
QY 261 AlaLeuLeuLeuAsnArgAsnAsnMetAlaAlaMetAspAspValProMetThrGln 280
Db 788 GCGTTGCTGCTGAATCGAAACAATATGGCAGCCATGGATGACGATGTTCTATGACACAG 847
QY 281 SerPheGlyCysGlySerTyrAlaLeu 289
Db 848 AGCTTTGGATGTGGAAGTTATGCTATT 874

RESULT 6
LOCUS AR228702 886 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 3 from patent US 6448046.
ACCESSION AR228702
VERSION AR228702.1 GI:27267789
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Unclassified.
1 (bases 1 to 886)
Donson,J., Dawson,W.O., Grantham,G.L., Turpen,T.H., Turpen,A.M.,
Garger,S.J. and Grill,L.K.
TITLE Recombinant animal viral nucleic acids
JOURNAL Patent: US 6448046-A 3 10-SEP-2002;
FEATURES
source
1..886
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores:
```



```

Db      488 ACTTTGTTTACTACACGCCAAATCTGCTGCGTGGCACTTATGGTACTCAITTCAGTCG 547
QY      181 ThrSerGluAlaAlaArgTyrLysPheIleGluGlnIleGlyLysArgValAspLys 200
Db      548 ACCTCTGAGCTCGGAGGTATAAATTTATTTAGACACAAATGGGAAGCGGTGACAAA 607
QY      201 ThrPheLeuProSerLeuAlaIleIleSerLeuGluAsnSerTyrPheSerAlaLeuSerLys 220
Db      608 ACCTTCCTACCAAGTTTAGCAATTTATAAGTTTGGAAAATAGTTGGTCTGCTCTCAAG 667
QY      221 GlnIleGlnIleAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuIleAsn 240
Db      668 CAAATTGAGATAGCGAGTACTATAATAATGGACAGCTTTGAAACTCCTGCTGTCTATAAAT 727
QY      241 AlaGlnAsnGlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnIle 260
Db      728 GCTCAAAACCAAGAGTCATGATACCAATGTTGATGCTGGAGTTGTAACCTCCAAATC 787
QY      261 AlaLeuLeuLeuAsnArgAsnAsnMetAlaAlaMetAspAspValProMetThrGln 280
Db      788 GCGTTGCTGCTGAATCGAAACAATATGGCAGCATGATGACGATGTTCTATGACACAG 847
QY      281 SerPheGlyCysGlySerTyrAlaLeu 289
Db      848 AGCTTTGGATGTGGAAGTTATGCTATT 874

RESULT 8
LOCUS      BD165859                      886 bp      DNA      linear      PAT 17-JAN-2003
DEFINITION Transgenic plant virus nucleic acid.
ACCESSION  BD165859
VERSION    BD165859.1 GI:27871671
KEYWORDS  JP 2002171976-A/2.
SOURCE    unidentified
ORGANISM  unclassified.
REFERENCE  1 (bases 1 to 886)
AUTHORS   Donson,J., Dawson,W.O., Grantham,G.L., Turpen,T.H., Turpen,A.M.,
          Gerger,S.J. and Grill,L.K.
TITLE     Transgenic plant virus nucleic acid
JOURNAL   Patent: JP 2002171976-A 2 18-JUN-2002;
          LARGE SCALE BIOLOGY CORP
COMMENT   OS Cucumis L. (chinese cucumber)
          PN JP 2002171976-A/2
          PD 18-JUN-2002
          PF 10-OCT-2001 JP 2001313207
          PR 01-AUG-1991 US 739143
          PI JOHN DONSON,WILLIAM O DAWSON,GEORGE L GRANTHAN,THOMAS H
          TURPEN,
          PI ANN MYERS TURPEN,STEPHEN J GERGER,LAURENCE K GRILL PC
          C12N15/09,C12N15/00
          CC Clone: alpha-trichosanthin
          FH Key Location/Qualifiers
          FT CDS Location/Qualifiers
              1..886
              /organism="unidentified"
              /mol_type="genomic DNA"
              /db_xref="taxon:32644"

FEATURES
source

ORIGIN

Alignment Scores:
Pred. No.:      2,85e-119      Length:      886
Score:          1413.00      Matches:      285
Percent Similarity: 99.31%      Conservative: 285
Best Local Similarity: 98.62%      Mismatches: 2
Query Match:    98.67%      Indels:      0
DB:             6      Gaps:      0

US-09-905-247-1 (1-289) x BD165859 (1-886)
QY      1 MetIleArgPheLeuValLeuSerLeuLeuIleLeuThrLeuPheLeuThrProAla 20

```

```

Db      8 ATGATCAGATTCCTTAGTCCTCTCTTTGCTAAATTCCTCACCCCTTCTCCATAACAATCTCGT 67
QY      21 ValGluGlyAspValSerPheArgLeuSerGlyValAlaThrSerSerSerTyrGlyValPhe 40
Db      68 GTGAGGGCGATGTTAGCTTCGTTTATCAGGTGCAACAGCAGTTCCTATGAGATTTC 127
QY      41 IleSerAsnLeuArgLysAlaLeuProAsnGluArgLysLeuTyrAspIleProLeuLeu 60
Db      128 ATTTCAAATCTGAGAAAAGCTCTTCCAAATGAAAGGAAACTGTACGATATCCCTCTGTTA 187
QY      61 ArgSerSerLeuProGlySerGlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAsp 80
Db      188 CGTTCCTCTCTCCAGTCTTCAACGCTACGATTCATCTCAAAATTAACGGCGAT 247
QY      81 GluThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGly 100
Db      248 GAAACCATTTAGTGCCCATAGACGTAAACGAGCTCTATATTATGGGATATCGCGTGGC 307
QY      101 AspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPheLys 120
Db      308 GATACATCCTATTTTTCACGAGGCTTCTGCAACAGAGCTTCGCAAGAGTGCATAATATGTA 367
QY      121 AspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAla 140
Db      368 GAGCGTATGCGAAAAGTTAGCTTTCATATTCTGGCAATTGCAAGGCTTCAAACTGCT 427
QY      141 AlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThr 160
Db      428 GCGGGCAAAATAAGGGAAAATATTCCGCTTGGACTCCAGCTTTGGACAGTCCCATACC 487
QY      161 ThrLeuPheTyrTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSer 180
Db      488 ACTTTGTTTACTACACGCCAAATTCGTGCTGCGCACTTATGCTACTCATTCAGTCG 547
QY      181 ThrSerGluAlaAlaArgTyrLysPheIleGluGlnIleGlyLysArgValAspLys 200
Db      548 AGCTCTGAGGCTCGGAGGTATAAATTTATGAGCAACAAATGGGAAGCGCTTGACAAA 607
QY      201 ThrPheLeuProSerLeuAlaIleIleSerLeuGluAsnSerTyrPheSerAlaLeuSerLys 220
Db      608 ACCTTCCTACCAAGTTTAGCAATTATAAGTTTGGAAAATAGTTGCTGCTCTCTCCAG 667
QY      221 GlnIleGlnIleAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuIleAsn 240
Db      668 CAAATTGAGATAGCGAGTACTATAATAATGGACAGCTTTGAAACTCCTCTGCTGTATAAAT 727
QY      241 AlaGlnAsnGlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnIle 260
Db      728 GCTCAAAACCAACGAGTCTATGATACCAATGTTGATGCTGGAGTTGTAACTCCCAATC 787
QY      261 AlaLeuLeuLeuAsnArgAsnAsnMetAlaAlaMetAspAspValProMetThrGln 280
Db      788 GCGTTGCTGCTGAATCGAAACAATATGGCAGCCATGATGACGATGTTCTCTATGACACAG 847
QY      281 SerPheGlyCysGlySerTyrAlaLeu 289
Db      848 AGCTTTGGATGTGGAAGTTATGCTATT 874

RESULT 9
TRHTCS

LOCUS      TRHTCS                      1227 bp      DNA      linear      PLN 16-JUL-1993
DEFINITION T.kirilowii alpha-trichosanthin (alpha-TCS) gene, complete cds.
ACCESSION  J05434
VERSION    J05434.1 GI:170534
KEYWORDS  alpha-trichosanthin; ribosome-inactivating protein; type I
          ribosome-inactivating protein.
SOURCE     Trichosanthin kirilowii (Mongolian snake-gourd)
ORGANISM  Trichosanthin kirilowii
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
          rosids; eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.
REFERENCE  1 (bases 1 to 1227)

```

AUTHORS Chow,T.P., Feldman,R.A., Lovett,M. and Piatlak,M.  
TITLE Isolation and DNA sequence of a gene encoding alpha-trichosanthin,  
a type I ribosome-inactivating protein  
JOURNAL J. Biol. Chem. 265 (15), 8670-8674 (1990)  
MEDLINE 90256790  
PUBMED 2341400  
COMMENT Original source text: T.kirilowii (strain Maximowicz) leaf DNA,  
clone pQ2D.  
Draft entry and computer-readable sequence for [1] kindly submitted  
by M.Piatlak, 15-MAR-1990, for release after publication.  
FEATURES  
source 1..1227  
/organism="Trichosanthus kirilowii"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:3677"  
25..30  
/note="putative"  
TATA\_signal 168..173  
/note="putative"  
TATA\_signal 238..243  
/note="putative"  
TATA\_signal 283..288  
/note="putative"  
340..1209  
/function="inactivates eukaryotic 60S ribosomal subunits"  
/codon\_start=1  
/product="trichosanthin"  
/protein\_id="AAA34206.1"  
/db\_xref="GI:170535"  
/translation="MIRPLVLSLILFLTPAVEGVDFSLGATSSSYGVFISNL  
RKALPNERKLYDIPLLRSLPGSQRYALIHNTVADETISVAIDVNVYIMGYRAGDT  
SYFNEASATAAKVKFDAMKVTLPYSGNVERLOTAAGKIRENIPLGLPALDSAIT  
ILFYNNASASALMVLQSTSEARAYKFIQQIGKRVDTFLPSLAIISLENSWSAL  
SKQIQIASTNNQGFETPVLLNAQQRVMTINVDGVVTSNIALLLNNNAAMDDVV  
PMTQSFQGGSYAI"  
340..408  
/note="alpha-trichosanthin signal peptide"  
409..1149  
/product="alpha-trichosanthin"  
ORIGIN  
Alignment Scores:  
Pred. No.: 4,388-119 Length: 1227  
Score: 1413.00 Matches: 285  
Percent Similarity: 99.31% Conservative: 2  
Best Local Similarity: 98.62% Mismatches: 2  
Query Match: 98.67% Indels: 0  
DB: 8 Gaps: 0  
US-09-905-247-1 (1-289) x TRHTCS (1-1227)  
QY 1 MetIleArgPheLeuValLeuSerLeuLeuLeuLeuThrLeuPheLeuThrProAla 20  
DB 340 ATGATCAGATTCTTAGTCTCTCTTTGCTTAATCTCACCTCTTCTCAACACTCTCTGCT 399  
QY 21 ValGluGlyAspValSerPheArgLeuSerGlyAlaThrSerSerSerGlyValPhe 40  
DB 400 GTGGAGGGCGATGTAGCTTCCTTTATCAGGTGCACAGAGATTCCTATGGAGTTTC 459  
QY 41 IleSerAsnLeuArgLysAlaLeuProAsnGluArgLysLeuThrAspLeuProLeuLeu 60  
DB 460 ATTTCAAATCTGAGAAAGCTCTTCCAAATGAAGGAAACTGTACGATATCTCTCTGTTA 519  
QY 61 ArgSerSerLeuProGlySerGlnArgGlyAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 80  
DB 520 CGTTCCTCTCTCCAGGTCTCAACGCTACGCAATGATCCATCTCAAAATACGCCGAT 579  
QY 81 GluThrIleSerValAlaIleAspValThrAsnValThrIleMetGlyTyrArgAlaGly 100  
DB 580 GAAACCATTTTCAGTGCCATAGACGTACGACGCTATATATATGGATATCGGCTGCG 639  
QY 101 AspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaIleTyrValPheLys 120

DB 640 GATACATCTCTATTTTTTCAACGAGGCTTCTGCAACAGAACTGCAAAATATGTATTTCAA 699  
QY 121 AspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAla 140  
DB 700 GAGCTATGCGAAAGATTACGCTTCCATATCTTGGCAATTTACGAAAGGCTTCAAACTGCT 759  
QY 141 AlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThr 160  
DB 760 CGCGGCAAAATAAGGGAATAATTTCCGCTTTGGACTCCAGCTTTGGACAGTGCATTACC 819  
QY 161 ThrLeuPheTyrTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSer 180  
DB 820 ACTTTGTTTACTACAAACGCCAATTTCTGCTGCGTCCGCACTTATGCTACTCATTCAGTCG 879  
QY 181 ThrSerGluAlaAlaArgTyrLysPheIleGluGlnIleGlyLysArgValAspLys 200  
DB 880 ACCTCTGAGGCTGCGAGGTATTAATTTATTGAGCAAAATTTGGGAAGCGGCTTGACAAA 939  
QY 201 ThrPheLeuProSerLeuAlaIleIleSerLeuGluAsnSerTyrSerAlaLeuSerLys 220  
DB 940 ACCTTCTCTACCAAGTTTAGCAATTTAAGTTTGGAAATAGTTGGTCTGCTCTCTTCCAA 999  
QY 221 GlnIleGlnIleAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuIleAsn 240  
DB 1000 CAATTCAGATAGGAGTACTAATTAATGACAGCTTTGAAACTCTCTGTTGCTTATAAT 1059  
QY 241 AlaGlnAsnGlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnIle 260  
DB 1060 GCTCAAAACCAACGAGTATGATCAATGTTGATGCTGAGTTGTAACCTCCCAACATC 1119  
QY 261 AlaLeuLeuLeuAsnArgAsnMetAlaAlaMetAspAspAspValProMetThrGln 280  
DB 1120 CGCTTGTCTGCTGAATCGAAACAATATGCGACCATGATGATGATGATGATGATGATGAT 1179  
QY 281 SerPheGlyCysGlySerTyrAlaLeu 289  
DB 1180 AGCTTTGGATGTGGAAGTTATGCTATT 1206  
RESULT 10  
S70176 1134 bp DNA linear PLN 23-SBP-1994  
LOCUS trichosanthin [Trichosanthus kirilowii, Maxim, Genomic, 1134 nt].  
DEFINITION S70176  
ACCESSION S70176.1 GI:547148  
VERSION S70176.1  
KEYWORDS Trichosanthus kirilowii (Mongolian snake-gourd)  
SOURCE Trichosanthus kirilowii  
ORGANISM Trichosanthus kirilowii  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.  
REFERENCE 1 (bases 1 to 1134)  
AUTHORS Zheng,H., Wang,B., Shaw,P. and Yeung,H.  
TITLE Cloning and DNA sequencing of the gene encoding trichosanthin  
JOURNAL Yi Chuan Xue Bao 21 (1), 42-51 (1994)  
MEDLINE 94271613  
PUBMED 8003348  
REMARK GenBank staff at the National Library of Medicine created this  
entry [NCBI gibbsg 148037] from the original journal article.  
This sequence comes from Fig. 4.  
FEATURES  
source location/Qualifiers  
1..1134  
/organism="Trichosanthus kirilowii"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:3677"  
244..1113  
/gene="trichosanthin, TCS"  
244..1113  
/gene="trichosanthin, TCS"  
/note="This sequence comes from Fig. 4; TCS"  
/codon\_start=1  
/product="trichosanthin"  
/protein\_id="AAB31048.1"  
/db\_xref="GI:547149"

Alignment Scores:				
Pred. No.:	1.12e-118	Length:	1134	
Score:	1408.00	Matches:	283	
Percent Similarity:	99.65%	Conservative:	5	
Best Local Similarity:	97.92%	Mismatches:	1	
Query Match:	98.32%	Indels:	0	
DB:	8	Gaps:	0	
US-09-905-247-1 (1-289) x S70176 (1-1134)				
QY	1	MetIleArgPheLeuValLeuSerLeuLeuLeuLeuThrLeuPheLeuThrProAla	20	
Db	244	ATGATCAGATCTTAGTCCTCTCTTTGCTAAATCTCACCTCTTCTTAAACAACCTCTGCT	303	
QY	21	ValGluGlyAspValSerPheArgLeuSerGlyAlaThrSerSerSerTyrGlyValPhe	40	
Db	304	GTGGAGGGCGATGTAGCTTCGTTTATCAGGTGCAACAGCAGTTCCTATGGAGTTTTC	363	
QY	41	IleSerAsnLeuArgIysAlaLeuProAsnGluArgLysLeuTyrAspIleProLeuLeu	60	
Db	364	ATTTCAAATCTGAGAAAGCTCTTCCAAATCAAGAGAACTGTACGATATGCCCTTGTGTA	423	
QY	61	ArgSerSerLeuProGlySerGlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAsp	80	
Db	424	CGTTCAGCTTCCAGGTTCTCAACGCTACGATGTGTCCATCTCACAAAATTACGCCGAT	483	
QY	81	GluThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGly	100	
Db	484	GAAAACCATTTTCAGTGGCCATAGACGTAAACGAGCGTCTATATTATGGCATATCGCGCTGGC	543	
QY	101	AspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaLysTyrValPheLys	120	
Db	544	GATACATCTCTATTTTTCACAGAGGCTTCGCAACAGAACTGCAAAATATGATTCCAAA	603	
QY	121	AspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAla	140	
Db	604	GAGCGTATCGCAAAAGTTAGCTTCCATNTCTGGCAATTACGAAAGCTTCAAACTGCT	663	
QY	141	AlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThr	160	
Db	664	GGGGGCAAAATAGGGAAATAATTCCGCTTGGACTCCAGCTTTGGACAGTGCCATTACC	723	
QY	161	ThrLeuPheTyrTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSer	180	
Db	724	ACTTTGTTTACTACAAACGCCAAATCTGCTGCGTCGGCACTTATGTGTACTCATTCAGTCG	783	
QY	181	ThrSerGluAlaAlaArgTyrLysPheIleGluGlnGlnIleGlyLysArgValAspLys	200	
Db	784	ACGCTCGAGCTCGGAGTATAAATTTATGACACAAATTCGGAGACGCCGTTGACAAA	843	
QY	201	ThrPheLeuProSerLeuAlaIleIleSerLeuGluAsnSerTrpSerAlaLeuSerLys	220	
Db	844	ACCTTCTCTACCAAGTTTAGCAATATATAAGTTTGGAAAAATAGTTGGTCTGCTCTCCAAG	903	
QY	221	GlnIleGlnIleAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuIleAsn	240	
Db	904	CAAAATTCAGATACGGAGTACTAATAATATGGACAGTTTGAACACTCCTGTGTGCTTTATAAT	963	
QY	241	AlaGlnAsnGlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnIle	260	
Db	964	GCTCAAAACCAACGAGTACGATAAACAATGTGTATGCTGGAGTTGTAACTCCACATC	1023	
QY	261	AlaLeuLeuLeuAsnArgAsnAsnMetAlaAlaMetAspAspValProMetThrGln	280	
Db	1024	GGGTGCTGCTGAATCGAAACGATATATGGAGCCATGATGATGACCATGTTCTCTATGACACAG	1083	



Db 367 GTGAGGCGGATGTTAGCTTCGGTTTATCAGGTGCAACAGCAGTTCCTATGGAGTTTC 426  
 Qy 41 IleserAsnLeuArgLysAlaLeuProAsnGluArgLysLeuTyrAspIleProLeuLeu 60  
 Db 427 ATTTCAATCTGAGAAAGCTCTTCCATATGAAAGGAACTGTACGATATCCCTCTCTGTA 486  
 Qy 61 ArgSerSerLeuProGlySerGlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAsp 80  
 Db 487 CGTTCCACTCTCCAGGTTCTCAACGCTTACCGATTCATGTCATCTCAAAATACGCCGAT 546  
 Qy 81 GluThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGly 100  
 Db 547 GAACACATTCAGTGGCCATAGAGCTAGACGCTCTATGTTATGGGATATCGCGTGT 606  
 Qy 101 AspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPheLys 120  
 Db 607 GATACATCCTATTATTTTCAACAGGCTTCTGCAACAGAGCTGCAAAATATGTATTCAA 666  
 Qy 121 AspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluAlaGluGlnThrAla 140  
 Db 667 GACGCTAAGCGAAAGATTACGCTTCCATATTCTGCAATTACGAAGGCTTCAATTGCT 726  
 Qy 141 AlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThr 160  
 Db 727 CGAGCAAAATAGGAAATATTCGCTTGAGCTTCCAGCTTTCGACAGTGCATTACC 786  
 Qy 161 ThrLeuPheTyrTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSer 180  
 Db 787 ACTTGTGTTTACTACAACGCCAATCTGCTGCTGCGCTGCGACCTATGCTACTCAATTCG 846  
 Qy 181 ThrSerGluAlaAlaArgTyrLysPheIleGluGlnIleGlyLysArgValAspLys 200  
 Db 847 ACCTCTGAGGCTGCGAGGTATATAATTATGAGCAACAAATTCGGAAGCGTGTGACAA 906  
 Qy 201 ThrPheLeuProSerLeuAlaIleSerLeuGluAsnSerTyrPheSerAlaLeuSerLys 220  
 Db 907 ACCTCTTACCAAGCTTGCAATTATAAGTTTGGAAATAGTTGCTGCTCTCTCCAAG 966  
 Qy 221 GlnIleGlnIleAlaSerThrAsnGlnPheGluSerProValValLeuIleAsn 240  
 Db 967 CAATTCAGATAGCAGTACTAATAATGACAGCTTGAACCTCTGTTGTGCTTATAAT 1026  
 Qy 241 AlaGlnAsnGlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnIle 260  
 Db 1027 GCTCAAAACCAAGAGTACAGATTAACCAATGTGTGCTGAGTGTGAACCTCCAATC 1086  
 Qy 261 AlaLeuLeuLeuAsnArgAsnAsnMetAlaAlaMetAspAspValProMetThrGln 280  
 Db 1087 GCGTTGCTGCTAAATAGAAACAATATGCGAGCTATTGATGACGATGTTCTATGGCACAG 1146  
 Qy 281 SerPheGlyCysGlySerTyrAlaLeu 289  
 Db 1147 AGCTTTGGATGTGAAGTTATGCTATT 1173

## RESULT 12

TKU25675 810 bp DNA linear PLN 20-JUL-1995  
 Trichosanthos kirilowii trichosanthin gene, partial cds.

DEFINITION  
 ACCESSION U25675  
 VERSION U25675.1 GI:847831

KEYWORDS  
 SOURCE Trichosanthos kirilowii (Mongolian snake-gourd)

ORGANISM  
 Trichosanthos kirilowii  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthos.

1 (bases 1 to 810)

REFERENCE  
 AUTHORS Bao, Y., Chu, R., Han, J., Zhang, H., Pan, N., Gu, X. and Chen, Z.  
 TITLE Cloning and sequencing of trichosanthin gene and its expression in  
 JOURNAL Escherichia coli and tobacco plant  
 Sci. China. B. 36 (No.6), 669-676 (1993)

2 (bases 1 to 810)

## AUTHORS

Direct Submission  
 TITLE Submitted (21-APR-1995) Lin Xu, Institute of Biophysics, Academia  
 JOURNAL Sinica, Dept. of Protein Engineering, 15 Datun Road, Chaoyang  
 District, Beijing 100101, Peoples Republic of China

## FEATURES

source Location/Qualifiers

1..810  
 /organism="Trichosanthos kirilowii"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:3677"  
 1..>810  
 /codon\_start=1  
 /label="JC1073"  
 /product="trichosanthin"  
 /protein\_id="AAA70096.1"  
 /db\_xref="GI:847832"  
 /translation="MIRFLVLSLLILTLFTTPAVGDSFRLSGARSSSYGVFISNL  
 RKALNERRRDIYDIPLESTPLGECERYALHLRYVADETISVAIDRVNRYIMGVKAGDT  
 FYFNEASATEAAKIVFKDMRWKVIPLPSGNRYERLQTAAGKIRENIPLIGLPALDSALT  
 TLFYFNANSAAALMWLIQSTSEAAKYKFEQQIGKRVKDTFIPSPAILISLNSWAL  
 SRQIQIASTNNGPESPVLINAEQRVTTITNDAGVVTNSIALLLRNNA"

## ORIGIN

Alignment Scores:  
 Pred. No.: 1.35e-105 Length: 810  
 Score: 1262.00 Matches: 256  
 Percent Similarity: 97.04% Conservative: 6  
 Best Local Similarity: 94.81% Mismatches: 8  
 Query Match: 88.13% Indels: 0  
 DB: 8 Gaps: 0

US-09-905-247-1 (1-289) x TKU25675 (1-810)

Qy 1 MetIleArgPheLeuValLeuSerLeuLeuIleLeuThrLeuPheLeuThrProAla 20  
 Db 1 ATGATCAGATTCTTAGTCTCTCTTTGCTATTCTCACCTCTTCTTAACAACCTCTGCT 60  
 Qy 21 ValGluGlyAspValSerPheArgLeuSerGlyAlaThrSerSerSerTyrGlyValPhe 40  
 Db 61 GTGAGGCGCATGTTAGCTTCGGTTATCAGGTGCAAGACAGTTCCTATGGAGTTTC 120  
 Qy 41 IleSerAsnLeuArgLysAlaLeuProAsnGluArgLysLeuTyrAspIleProLeuLeu 60  
 Db 121 ATTTCAATCTGAGAAAGCTCTTGGAAATGAAGAGACTATATGATATCCCTCTGTA 180  
 Qy 61 ArgSerSerLeuProGlySerGlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAsp 80  
 Db 181 CGTTCCACTCTTCCAGGTTGTGAACGCTACGATTCATTCATCTCAGAAATACCGCAT 240  
 Qy 81 GluThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGly 100  
 Db 241 GAACCAATTCAGTGGCCATAGACGTAAGGAACGCTCTATATTATGGGATATCGCGTGC 300  
 Qy 101 AspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPheLys 120  
 Db 301 GATACATTCATTTTTTCAACAGGCTTCTGCAACAGAGCTGCAAAATATGATTTCAA 360  
 Qy 121 AspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAla 140  
 Db 361 GACGTTATGCGAAAGTTACGCTTCCATATTCTGCAATTCAGAAAGCTTCAAACTGCT 420  
 Qy 141 AlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThr 160  
 Db 421 GCAGGCAAAATCAGGAAATATTCGCTTGGACTCCAGCTTTGGACAGTCCCATACC 480  
 Qy 161 ThrLeuPheTyrTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSer 180  
 Db 481 ACTTGTGTTTACTACAACGCCAATTTCTGCTGCTGCGACTATGCTACTCATTCAGTCG 540  
 Qy 181 ThrSerGluAlaAlaArgTyrLysPheIleGluGlnIleGlyLysArgValAspLys 200  
 Db 541 ACCTCTGAGGCTCGGAGGTATAAATTTATTGAGCAACAAATTCGGAAGCGCGTTGACAA 600



```
QY 201 ThrPheLeuProSerLeuAlaIleSerLeuGluAsnSerTyrPheLeuSerLys 220
Db 601 ACCTTCATACCAAGTTTGGCAATTATAAGTTTGGAAAATAGTTGGTCTCTCTCAGG 660
QY 221 GlnIleGlnIleAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuIleAsn 240
Db 661 CAAATTACAGATACGAGTACTAATAATGACAGTTTGAAGTCTCTGTGTGCTTATAAAT 720
QY 241 AlaGlnAsnGlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnIle 260
Db 721 GCTGAACCAACAGGATCAGATCACAATGTTGATGCTGGAGTTGTAACCTCCAAATC 780
QY 261 AlaLeuLeuLeuAsnArgAsnMetAla 270
Db 781 GCGTTGCTGCTGAATAGAAAACAATATGGCA 810

RESULT 13
AY082348
LOCUS
DEFINITION
Trichosanthin kirilowii trichosanthin precursor (TCS) gene, partial
cds.
ACCESSION
AY082348
VERSION
AY082348.1 GI:20977594
KEYWORDS
Trichosanthin kirilowii (Mongolian snake-gourd)
SOURCE
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthin.
1. (bases 1 to 1433)
AUTHORS
Yuan,H., Wang,L., Wang,X., Liu,T., An,C. and Chen,Z.
TITLE
Trichosanthin kirilowii trichosanthin precursor (TCS) gene
REFERENCE
1. (bases 1 to 1433)
JOURNAL
Wang,L., Yuan,H., Liu,T., An,C. and Chen,Z.
Direct Submission
Submitted (05-MAR-2002) National Laboratory of Protein Engineering
and Plant Genetic Engineering, College of Life Sciences, Peking
University, Beijing 100871, China
FEATURES
Location/Qualifiers
source
1..1433
/organism="Trichosanthin kirilowii"
/mol_type="genomic DNA"
/db_xref="taxon:3677"
1..>1433
/gene="TCS"
1..598
/gene="TCS"
misc_feature
597
/gene="TCS"
/note="alternative weaker transcriptional start site"
599..>1433
/gene="TCS"
/product="trichosanthin precursor"
599..623
/gene="TCS"
600
/gene="TCS"
600
/gene="TCS"
/note="alternative weaker transcriptional start site"
624..>1433
/gene="TCS"
/note="ribosome inactivating protein"
/codon_start=1
/product="trichosanthin precursor"
/protein_id="AA022782.1"
/db_xref="GI:20977595"
translation="MIRFLVLSLILFLTPAPAVEGDVSPRLSGATSSSYGVFISNL
RKALPYEKKYDIPIILRSSLSQSORIALIYLTNVADETISVAIDVTNVIYMGYAGDI
SYFFNEASATEAAKVFQDAKRVTLPYSGNVERLQIAAGKIRENIPGLPALDSAIT
TLFYVNANSAALLVLTQSTAEAAKYKFIQQIGKRVDTKFLPSLAISIENSWSL
SKQIQIATNNQFESPVLINAQNRVTIANVAGVVISNIAFLNLRNMA"
624..692
/gene="TCS"

sig_peptide
```

```
mat_peptide 693..1433
/gene="TCS"
/product="trichosanthin"
```

## ORIGIN

```
Alignment Scores:
Pred. No.: 1-88e-104 Length: 1433
Score: 1253.00 Matches: 256
Percent Similarity: 96.67% Conservative: 5
Best Local Similarity: 94.81% Mismatches: 9
Query Match: 87.50% Indels: 0
DB: Gaps: 0
```

US-09-905-247-1 (1-289) x AY082348 (1-1433)

```
QY 1 MetIleArgPheLeuValLeuSerLeuLeuIleLeuThrLeuPheLeuThrProAla 20
Db 624 ATGATCAGATTCTTAGTCTCTCTTTGCTAAATCTCACCTCTCTCTTAACAGCTCTGCT 683
QY 21 ValGluGlyAspValSerPheArgLeuSerGlyAlaThrSerSerSerTyrGlyValPhe 40
Db 684 GTCGAGGGCGATGTTAGCTTCCTGTTATCTGCTGCAACAGCAGTTCTTATGAGTTTTC 743
QY 41 IleSerAsnLeuArgGlyAlaLeuProAsnGluArgLysLeuTyrAspIleProLeuLeu 60
Db 744 ATTTCAATCTGAGAAAGCTCTTCATATGAAGAAGAACTGTACGATATCCCTCTGTTA 803
QY 61 ArgSerSerLeuProGlySerGlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAsp 80
Db 804 CGTTCCTCTCTTTCAGGTTCTCAACGCTACGCAATTAATCTACTCACAATATACGCCGAT 863
QY 81 GluThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGly 100
Db 864 GAACCATTTTCAGTGGCCATACACGTAACGACGCTATATATATGGGATATCGCGTGT 923
QY 101 AspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaAlaIysTyrValPheLys 120
Db 924 GATATATCTTATTTTTCACGAGGCTCTCGCAACAGAGCTGCAAAATATATATTCAAA 983
QY 121 AspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAla 140
Db 984 GACGCTAAGCGAAAAGTTACGCTTCCATATCTCGAAATACGAAGGCTTCAAATTGCT 1043
QY 141 AlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThr 160
Db 1044 GCAGCAAAATTAAGGAAAAATATTCGCTTGGACTCCCGAGCTTTGGACAGTGCCTATACC 1103
QY 161 ThrLeuPheTyrTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSer 180
Db 1104 ACTTTGTTTACTACACGCCAAATTCGCTGCTCGGCACTTCTCTACTCANTCAGTCC 1163
QY 181 ThrSerGluAlaAlaArgTyrLysPheIleGluGlnGlnIleGlyLysArgValAspLys 200
Db 1164 ACGGCTGAGGCTGCGAGGTATAAAATTTATTGAGCAACAAATTTGGAGCGGTGTTGACAAA 1223
QY 201 ThrPheLeuProSerLeuAlaIleSerLeuGluAsnSerTyrPheAlaLeuSerLys 220
Db 1224 ACCTTCTTACCAGCTAGCAATTAATAAGTTTGGAAAAATAGTTGCTCTGCTCTTCCAAG 1283
QY 221 GlnIleGlnIleAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuIleAsn 240
Db 1284 CAATTCAGATAGCGAGTACTAATAATGACAGCTTTGAAAGTCTCTGTTGCTTATAAAT 1343
QY 241 AlaGlnAsnGlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnIle 260
Db 1344 GCTCAAAACCAACAGATCAGCATAGCCAATGTTAATGCTGGAGTTGTAACTCCCAACATC 1403
QY 261 AlaLeuLeuLeuAsnArgAsnMetAla 270
Db 1404 GCGTTTCTGCTAAATAGAAAACAATATGGCA 1433
```

RESULT 14  
AB039324

```

LOCUS      AB039324                      741 bp    DNA          linear    PLN 19-JUN-2002
DEFINITION Trichosanthes sp. Bac Kan 8-98 rfb gene for Trichobakin, partial
            cds.
ACCESSION  AB039324
VERSION    AB039324.1  GI:7242889
SOURCE     Trichosanthes sp. Bac Kan 8-98
ORGANISM   Trichosanthes sp. Bac Kan 8-98
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.
REFERENCE  1
            Van Chi, P., Quoc Trung, H., Thuy Ha, N., Chung, W.I. and Binh, L.T.
            Characterization of trichobakin, a type I ribosome-inactivating
            protein from Trichosanthes sp. Bac Kan 8-98
            Biotechnol. Appl. Biochem. 34 (2), 85-92 (2001)
            1172374
REFERENCE  2  (bases 1 to 741)
            Van Chi, P., Thuy Ha, N. and Binh, L.T.
            Direct Submission
            Submitted (28-FEB-2000) Phan Van Chi, Institute of Biotechnology,
            Protein Biochemistry, Hoang Quoc Viet, Hanoi 84-04, Vietnam
            (E-mail: pchi@netnam.vn, Tel: 7561903)
FEATURES   Location/Qualifiers
            1..741
             /organism="Trichosanthes sp. Bac Kan 8-98"
             /mol_type="genomic DNA"
             /db_xref="taxon:118182"
             /clone="pQV-20"
             /sex="female"
             /cell_type="Leaf"
             /dev_stage="early"
             /note="isolated from Province Bac Kan in the North of
             Vietnam august 1998-vector: pET-21d(+)"
            1..741
             /gene="TBK"
             /gene="TBK"
             /standard_name="type 1 ribosome inactivating protein"
             /codon_start=1
             /product="Trichobakin"
             /protein_id="BAA92530.1"
             /db_xref="GI:7242890"
             /translation="DVSFRLSGATSSSYGFISNLKALPYERKLYDIPLLRSTLPGS
             QRYALIHITNVADETISVALDVTNVVYMGAGDTSYFNEASATEAAKYVFKDAKRK
             VILPSGNYERLQIAAGKIRENIPGLPALSADITLFLVYNANSAALMWLQSTSE
             AARYKFIEQIGKRVKDKTFLPSLAISLNSWSALSKQIQIASTNNGQFETPVVLINA
             QNQRVTITNDAGVVTSTNIALLPNRNMA"
ORIGIN
Alignment Scores:
Pred. No.:      3,41e-98      Length:      741
Score:          1180.00      Matches:     239
Percent Similarity: 97.98%      Conservative: 3
Best Local Similarity: 96.76%      Mismatches:  5
Query Match:     82.40%      Indels:      0
DB:              8            Gaps:         0

US-09-905-247-1 (1-289) x AB039324 (1-741)

QY 24 AspValSerPheArgLeuSerGlyAlaThrSerSerSerTyrGlyValPheLeuSerAsn 43
Db 1 GATGTTAGCTTCGTTTATCAGGTGCAACAAGCAGTTCCTATGAGAGTTCATTTTCAAAAT 60
QY 44 LeuArgLysAlaLeuProAsnGluArgLysLeuTyrAspIleProLeuLeuArgSerSer 63
Db 61 CTGAGAAAGCTCTTCATATGAAAGAAACTGTACGATATCCCTCTGTTACGTTCCACT 120
QY 64 LeuProGlySerGlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAspGluThrIle 83
Db 121 CTTCCAGGTTCTCAACGCTACGATTGATCCATCTCACAAATTCGCCGATGAACCAATT 180
QY 84 SerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGlyAspThrSer 103

```

```

Db 181 TCAGTGGCCATAGACGTAACGAACGCTCTATGTTATGGGATATCGCGCTGGTGATACATCC 240
QY 104 TyrPhePheAsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPheLysAspAlaMet 123
Db 241 TATTTTTCACACGAGGCTTCTGCACAGAAAGCTGCCAAATATGTTTCAAGACGCTAAG 300
QY 124 ArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAlaAlaGlyLys 143
Db 301 CGAAAGTTACGCTTCATATCTGGCAATTACGAAAGCTTCAAAATGCTGCAGGCAAA 360
QY 144 IleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThrThrIlePhe 163
Db 361 ATAAGGGGAAATATTCGCTTGGACTCCAGCTTGGACAGTGGCAATTACCACCTTTGTTT 420
QY 164 TyrTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSerThrSerGlu 183
Db 421 TACTACACGCAATATCTGCTGCTGCGCACTTATGGTACTCATTCAGTCAGCGTCTGAG 480
QY 184 AlaAlaArgTyrLysPheIleGluGlnGlnIleGlyLysArgValAspLysThrPheLeu 203
Db 481 GCTCGAGGTATATAAATTTATTGAGCAACAAATTTGGAGCGTGTGTGACAAACCTTCCTA 540
QY 204 ProSerLeuAlaIleIleSerLeuGluAsnSerTrpSerAlaLeuSerLysGlnIleGln 223
Db 541 CCAAGTCTTGCATTTATAAGTTTGGAAATAGTTGGTCTGCTCTCTCCAGCAAAATTCAG 600
QY 224 IleAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuIleAsnAlaGlnAsn 243
Db 601 ATAGCGAGTACTAATATGACAGCTTGGAACTCTCTGTTGGCTTATAAATGCTCAAAAC 660
QY 244 GlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnIleAlaLeuLeu 263
Db 661 CAACGAGTCACGATAACCAATGTTGATGCTGGAGTTGTAACCTCCACATCGCGTGTGCTG 720
QY 264 LeuAsnArgAsnAsnMetAla 270
Db 721 CCAATAGAAACAAATATGGCA 741

RESULT 15
LOCUS      BYNBD1                      1084 bp    mRNA          linear    PLN 21-NOV-2001
DEFINITION Bryonia dioica ribosome-inactivating protein mRNA, complete cds.
ACCESSION  L42298
VERSION    L42298.1  GI:17028175
KEYWORDS   bryodin 1; ribosome inactivating protein.
SOURCE     Bryonia dioica (red bryony)
ORGANISM   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Cucurbitales; Cucurbitaceae; Bryonia.
REFERENCE  1  (bases 1 to 1084)
            Gawlak, S.L., Neubauer, M.G. and Siegall, C.B.
            Cloning, expression, and structure-function analysis of bryodin 1,
            a ribosome-inactivating protein from the plant Bryonia dioica
            Unpublished
            2  (bases 1 to 1084)
            Neubauer, M.G.
            Direct Submission
            Submitted (01-DEC-1996) M.G. Neubauer, Molecular Biology,
            Bristol-Myers Squibb Pharm. Res. Inst., 3005 1st Ave., Seattle, WA,
            USA, 98121 neubauer@bms.com
FEATURES   Location/Qualifiers
            1..1084
             /organism="Bryonia dioica"
             /mol_type="mRNA"
             /db_xref="taxon:3652"
             44..916
             /note="precursor
             putative"
             /codon_start=1
             /product="ribosome-inactivating protein"
             /protein_id="AAL34334.1"
CDS

```





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 14, 2004, 20:00:06 ; Search time 87 Seconds  
(without alignments)  
1843.458 Million cell updates/sec

Title: US-09-905-247-1

Perfect score: 1432

Sequence: 1 MRFLVLSLLILTLFTTPA.....AAMDDVPMTQSGCGSYAL 289

Scoring table:

	Xgapop 10.0	Xgapext 0.5
Ygapop 10.0	0.5	
Fgapop 6.0	0.5	
Delop 6.0	7.0	

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

```
-MODEL=frame_plus_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US0905247/runat_10092004_103934_4747/app_query.fasta_1.455
-DB=Issued Patents NA -OFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCTL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pcp -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0905247@cgn2_1 1 105 @runat_10092004_103934_4747 -NCPU=6 -ICPU=3
-NO MAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database : Issued Patents NA:

	1:	2:	3:	4:	5:	6:
/cgn2_6/ptodata/2/ina/5A.COMB.seq:						
/cgn2_6/ptodata/2/ina/5B.COMB.seq:						
/cgn2_6/ptodata/2/ina/6A.COMB.seq:						
/cgn2_6/ptodata/2/ina/6B.COMB.seq:						
/cgn2_6/ptodata/2/ina/6C.COMB.seq:						
/cgn2_6/ptodata/2/ina/backfiles1.seq:						

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1413	98.7	886	1	US-07-923-692C-3
2	1413	98.7	886	1	US-08-184-237-3
3	1413	98.7	886	2	US-08-482-920-3
4	1413	98.7	886	3	US-08-484-341-3
5	1413	98.7	886	3	US-08-483-502-3
6	1413	98.7	886	4	US-09-726-851A-3
7	1164.5	81.3	1094	1	US-08-245-754A-1
8	1164.5	81.3	1094	2	US-08-537-731-1
9	1074	75.0	1499	3	US-08-902-486-14
10	1064	74.3	758	3	US-08-902-486-6
11	728	50.8	913	1	US-07-901-707-13
12	728	50.8	913	1	US-07-988-430-13

13	728	50.8	913	1	US-08-425-336-13	Sequence 13, Appl
14	728	50.8	913	1	US-08-488-113B-13	Sequence 13, Appl
15	728	50.8	913	1	US-08-477-484B-13	Sequence 13, Appl
16	728	50.8	913	2	US-08-646-360-13	Sequence 13, Appl
17	728	50.8	913	3	US-08-839-765-13	Sequence 13, Appl
18	728	50.8	913	3	US-09-136-389-13	Sequence 13, Appl
19	728	50.8	913	4	US-09-610-838-13	Sequence 13, Appl
20	728	50.8	913	4	US-09-711-485-13	Sequence 13, Appl
21	728	50.8	913	5	PCT-US92-09487-13	Sequence 13, Appl
22	537.5	37.5	962	1	US-08-324-301-14	Sequence 14, Appl
23	427	29.8	1855	4	US-09-403-752A-31	Sequence 31, Appl
24	427	29.8	1879	4	US-09-147-208-23	Sequence 23, Appl
25	427	29.8	1879	4	US-09-550-117A-23	Sequence 23, Appl
26	425	29.7	1605	2	US-08-356-786-9	Sequence 9, Appl
27	423	29.5	1698	4	US-09-785-921A-1	Sequence 1, Appl
28	423	29.5	1805	4	US-09-785-921A-10	Sequence 10, Appl
29	423	29.5	1855	4	US-09-147-208-33	Sequence 33, Appl
30	423	29.5	1855	4	US-09-147-208-40	Sequence 40, Appl
31	423	29.5	1855	4	US-09-147-208-47	Sequence 47, Appl
32	423	29.5	1855	4	US-09-147-208-54	Sequence 54, Appl
33	423	29.5	1855	4	US-09-550-117A-33	Sequence 3, Appl
34	423	29.5	1855	4	US-09-550-117A-40	Sequence 40, Appl
35	423	29.5	1855	4	US-09-550-117A-47	Sequence 47, Appl
36	423	29.5	1855	4	US-09-550-117A-54	Sequence 54, Appl
37	423	29.5	1855	4	US-09-403-752A-3	Sequence 3, Appl
38	423	29.5	1855	4	US-09-403-752A-5	Sequence 5, Appl
39	423	29.5	1855	4	US-09-403-752A-7	Sequence 7, Appl
40	423	29.5	1855	4	US-09-403-752A-9	Sequence 9, Appl
41	423	29.5	1855	4	US-09-403-752A-13	Sequence 13, Appl
42	423	29.5	1855	4	US-09-403-752A-15	Sequence 15, Appl
43	423	29.5	1855	4	US-09-403-752A-17	Sequence 17, Appl
44	423	29.5	1855	4	US-09-403-752A-19	Sequence 19, Appl
45	423	29.5	1855	4	US-09-403-752A-21	Sequence 21, Appl

#### ALIGNMENTS

##### RESULT 1

US-07-923-692C-3

; Sequence 3, Application US/07923692C

; Patent No. 5316931

; GENERAL INFORMATION:

; APPLICANT: Donson, Jon

; APPLICANT: Dawson, William O.

; APPLICANT: Grantham, George L.

; APPLICANT: Turpen, Thomas H.

; APPLICANT: Turpen, Ann Myers

; APPLICANT: Garger, Stephen J.

; APPLICANT: Grill, Laurence K.

; TITLE OF INVENTION: RECOMBINANT PLANT VIRAL NUCLEIC ACIDS

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Limbach & Limbach

; STREET: 2001 Ferry Building

; CITY: San Francisco

; STATE: CAL

; ZIP: 94111

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/923,692C

; FILING DATE: 31-JUL-1992

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 600,244

; FILING DATE: 22-OCT-1990

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 641,617

; FILING DATE: 16-JAN-1991

; PRIOR APPLICATION DATA:

101 AspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPheLys 120  
 308 GATACATCTATTTTCAACGAGGCTTCTCAACAGAGAGCTGCAAAATATGTATTCAA 367  
 121 AspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAla 140  
 368 GACGCTATGCGAAAGTACGCTTCCATATCTGGCAATACGAAAGGCTTCAACTGCT 427  
 141 AlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThr 160  
 428 GCGGGCAAAATAGGGAAATATTCGCTTGGACTCCAGCTTGGACAGTGCATTACC 487  
 161 ThrLeuPheTyrTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSer 180  
 488 ACTTTGTTTACTACACGCAATTCGCTGCGTGGCACTATGGTACTCAITTCAGTCG 547  
 181 ThrSerGluAlaAlaArgTyrLysPheIleGluGlnGlnIleGlyLysArgValAspLys 200  
 548 ACCTCTGAGGCTGCGAGGTATTAATTTATGAGCAACAATTTGGAAAGCGGTTGACAAA 607  
 201 ThrPheLeuProSerLeuAlaIleIleSerLeuGluAsnSerTyrSerAlaLeuSerLys 220  
 608 ACCTTCTTACCAAGTTTAGCAATTTAAGTTTGGAAATAGTTGCTCTCTCTTCCAAAG 667  
 221 GlnIleGlnIleAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuIleAsn 240  
 668 CAAATTCAGATAGCGAGTACTAATTAATGAGCAGATTTGAAATCTCTGTGTCTTATAAAT 727  
 241 AlaGlnAsnGlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnIle 260  
 728 GCTCAAAACCAACGAGTCATGATTAACCAATCTTGTGTCTGAGTGTGAACCTCCCAACATC 787  
 261 AlaLeuLeuLeuAsnArgAsnAsnMetAlaAlaMetAspAspValProMetThrGln 280  
 788 GCGTTGCTGCTGAATCGAAACAATATGCGACCATGATGATGATGATGATGATGATGAT 847  
 281 SerPheGlyCysGlySerTyrAlaLeu 289  
 848 AGCTTTGGATGTGGAAGTTATGCTATT 874

RESULT 2  
 US-08-184-237-3  
 ; Sequence 3, Application US/08184237  
 ; Patent No. 5589367  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Donson, Jon  
 ; APPLICANT: Dawson, William O.  
 ; APPLICANT: Grantham, George L.  
 ; APPLICANT: Turpen, Thomas H.  
 ; APPLICANT: Turpen, Ann Myers  
 ; APPLICANT: Garger, Stephen J.  
 ; APPLICANT: Grill, Laurence K.  
 ; TITLE OF INVENTION: RECOMBINANT PLANT VIRAL NUCLEIC ACIDS  
 ; NUMBER OF SEQUENCES: 11  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Limbach & Limbach  
 ; STREET: 2001 Ferry Building  
 ; CITY: San Francisco  
 ; STATE: CAL  
 ; ZIP: 94111  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent in Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/184,237  
 ; FILING DATE:  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 923,692  
 ; FILING DATE: 31-JUL-1992  
 ; PRIOR APPLICATION DATA:

Alignment Scores: 1.06e-166 Length: 886  
 Pred. No.: 285  
 Score: 1413.00  
 Percent Similarity: 99.31%  
 Best Local Similarity: 98.62%  
 Query Match: 98.67%  
 DB: 1  
 Gaps: 0

US-09-905-247-1 (1-289) x US-07-923-692C-3 (1-886)  
 QY 1 MetIleArgPheLeuValLeuSerLeuLeuLeuThrLeuPheLeuThrProAla 20  
 Db 8 ATGATCAGATCTTAGTCTCTCTTTGCTTAATCTCACCTCTTCTCAACACTCTGCT 67  
 QY 21 ValGluGlyAspValSerPheArgLeuSerGlyAlaThrSerSerTyrGlyValPhe 40  
 Db 68 GTGAGGGCGATGTAGCTTCGCTTTATCAGTGCACACAGCAGTTCCTATGGAGTTTC 127  
 QY 41 IleSerAsnLeuArgLysAlaLeuProAsnGluArgLysLeuTyrAspIleProLeuLeu 60  
 Db 128 ATTTCAAAATCGAAAGTCTTCCAAATGAAAGAACTGTACGATATCCCTCTGTTA 187  
 QY 61 ArgSerSerLeuProGlySerGlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAsp 80  
 Db 188 CGTTCCTCTCTTCCAGTTCTCAACGCTACGATTCATTCATTCACAAATACGCCGAT 247  
 QY 81 GluThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGly 100  
 Db 248 GAAACCAATTCAGTGGCGCATAGACGTAACGAACGCTATATATGATGATATCGCGCTGC 307

APPLICATION NUMBER: US 310,881  
 FILING DATE: 17-FEB-1989  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 160,766  
 FILING DATE: 26-FEB-1988  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 160,771  
 FILING DATE: 26-FEB-1988  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 347,637  
 FILING DATE: 05-MAY-1989  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 363,138  
 FILING DATE: 08-JUN-1989  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 219,279  
 FILING DATE: 15-JUL-1988  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Halluin, Albert P.  
 REGISTRATION NUMBER: 28,957  
 REFERENCE/DOCKET NUMBER: BIOG-20121  
 REFERENCE/DOCKET NUMBER: USA  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 415-433-4150  
 TELEFAX: 415-433-8716  
 INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 886 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 ORIGINAL SOURCE:  
 ORGANISM: Chinese cucumber  
 IMMEDIATE SOURCE:  
 CLONE: alpha-trichosanthin  
 FEATURE:  
 NAME/KEY: CDS (B) LOCATION: 8..877  
 LOCATION: 8..877



Wed Sep 15 10:32:02 2004

```

PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 184,237
  FILING DATE: 19-JAN-1994
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 600,244
  FILING DATE: 22-OCT-1990
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 641,617
  FILING DATE: 16-JAN-1991
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 310,881
  FILING DATE: 17-FEB-1989
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 160,766
  FILING DATE: 26-FEB-1988
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 160,771
  FILING DATE: 26-FEB-1988
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 347,637
  FILING DATE: 05-MAY-1989
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 363,138
  FILING DATE: 08-JUN-1989
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 219,279
  FILING DATE: 15-JUL-1988
ATTORNEY/AGENT INFORMATION:
  NAME: Halluin, Albert P.
REGISTRATION NUMBER: 28,957
REFERENCE/DOCKET NUMBER: 8129-112
TELECOMMUNICATION INFORMATION:
  TELEPHONE: 415-854-3660
  TELEFAX: 415-854-3694
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
  LENGTH: 886 base pairs
  TYPE: nucleic acid
  STRANDEDNESS: single
  TOPOLOGY: linear
  MOLECULE TYPE: DNA (genomic)
  HYPOTHETICAL: NO
  ANTI-SENSE: NO
  ORIGINAL SOURCE:
  ORGANISM: Chinese cucumber
  IMMEDIATE SOURCE:
  CLONE: alpha-trichosanthin
  FEATURE:
  NAME/KEY: CDS (B) LOCATION: 8..877
  LOCATION: 8..877
US-08-482-920-3

Alignment Scores:
Pred. No.: 1,06e-166 Length: 886
Score: 1413.00 Matches: 285
Percent Similarity: 99.31% Conservative: 2
Best Local Similarity: 98.62% Mismatches: 2
Query Match: 98.67% Indels: 0
DB: 2 Gaps: 0

US-09-905-247-1 (1-289) x US-08-482-920-3 (1-886)
QY 1 MetIleArgPheLeuValLeuSerLeuLeuLeuThrLeuPheLeuThrProAla 20
Db 8 ATGATCAGATTCTTAGTCTCTCTTTGCTAAATTTCTCACCCCTCTTCTTAACTCCGCT 67
QY 21 ValGluGlyAspValSerPheArgLeuSerGlyValAlaThrSerSerSerTyrGlyValPhe 40
Db 68 GTGAGGGCGGATGTAGCTTCCTCCGTTTATCAGGTGCACACAGCAGTCTCTATGAGTTTC 127
QY 41 IleSerAsnLeuArgLysAlaLeuProAsnGluArgLysLeuTyrAspIleProLeuLeu 60
Db 128 ATTTCAAATCTGAGAAAAGCTCTTCCAAATGMAAGGAAACTGTACGATATCCCTCTGT 187

61 ArgSerSerLeuProGlySerGlnArgTyrAlaLeuLeuIleHisLeuThrAsnTyrAlaAsp 80
188 CGTTCTCTCTCTCCAGGTTCTCAACGCTACGATTCATCATCTCAAAATACGCCGAT 247
QY 81 GluThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGly 100
248 GAAACCATTTTCAGTGGCCATAGACGTAAACGAACGCTCTATATTAGGATATCCGCTGGC 307
QY 101 AspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPheLys 120
308 GATACATCTCTATTTTTCACAGAGGCTTCTGCAACAGAGCTGCAAAATATGTTATTTCAA 367
QY 121 AspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAla 140
368 GACGCTATCGAAAAAGTTACGCTTCCATATTTCTGGCAATTTACGAAAGGCTTCAAACTGCT 427
QY 141 AlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThr 160
428 GCGGGCAAAATAAGGGAAATATTTCCGCTTGGACTCCAGCTTCCAGCTTGGACAGTGCATTACC 487
QY 161 ThrLeuPheTyrTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSer 180
488 ACTTTGTTTACTACACGCGCAATTTCTGCTGCGTCCGCACTTATGTTACTCATTCAGTCG 547
QY 181 ThrSerGluAlaAlaArgTyrLysPheIleGluGlnIleGlyLysArgValAspLys 200
548 ACCTTCTACCAAGTTTAGCAATTTAAGTTTGGAAATAAGTTGGTGTCTCTCTCTCCCAAG 607
QY 201 ThrPheLeuProSerLeuAlaIleSerLeuGluAsnSerTyrPheSerAlaLeuSerLys 220
608 ACCTTCTACCAAGTTTAGCAATTTAAGTTTGGAAATAAGTTGGTGTCTCTCTCTCCCAAG 667
QY 221 GlnIleGlnIleAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuIleAsn 240
668 CAAATTCAGATAGCGAGTACTAATAATGGACAGATTGAACTCTCTGTTGCTGTATAAAT 727
QY 241 AlaGlnAsnGlnArgValThrIleThrAsnValAlaGlyValValThrSerAsnIle 260
728 GCTCAAAACCAACGAGTCAATGATACCAATGTTGATGCTGGAGTTGTAACCTTCAACATC 787
QY 261 AlaLeuLeuLeuAsnArgAsnAsnMetAlaAlaMetAspAspValProMetThrGln 280
788 GCGTTGCTGCTGAATCGAAACAAATATGCGCCATGATGACGATGATGATGATGATGATGAT 847
QY 281 SerPheGlyCysGlySerTyrAlaLeu 289
848 ACCTTTGGATGTGGAAGTTATGCTATT 874
Db

RESULT 4
US-08-484-341-3
; Sequence 3, Application US/08484341
; GENERAL INFORMATION:
; APPLICANT: Donson, Jon
; Dawson, William O.
; Grantham, George L.
; Turpen, Thomas H.
; Turpen, Ann Myers
; Garger, Stephen J.
; Grill, Laurence K.
; TITLE OF INVENTION: RECOMBINANT PLANT VIRAL NUCLEIC ACIDS
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESS: Limbach & Limbach
; STREET: 2001 Ferry Building
; CITY: San Francisco
; STATE: CAL
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25

```



<b>CURRENT APPLICATION DATA:</b>					
APPLICATION NUMBER: US/08/484,341					
FILING DATE: 07-Jun-1995					
CLASSIFICATION: <Unknown>					
<b>PRIOR APPLICATION DATA:</b>					
APPLICATION NUMBER: 08/184,237					
FILING DATE: <Unknown>					
APPLICATION NUMBER: US 600,244					
FILING DATE: 22-OCT-1990					
APPLICATION NUMBER: US 641,617					
FILING DATE: 16-JAN-1991					
APPLICATION NUMBER: US 310,881					
FILING DATE: 17-FEB-1989					
APPLICATION NUMBER: US 160,766					
FILING DATE: 26-FEB-1988					
APPLICATION NUMBER: US 160,771					
FILING DATE: 26-FEB-1988					
APPLICATION NUMBER: US 347,637					
FILING DATE: 05-MAY-1989					
APPLICATION NUMBER: US 363,138					
FILING DATE: 08-JUN-1989					
APPLICATION NUMBER: US 219,279					
FILING DATE: 15-JUL-1988					
<b>ATTORNEY/AGENT INFORMATION:</b>					
NAME: Halluin, Albert P.					
REGISTRATION NUMBER: 28,957					
REFERENCE/DOCKET NUMBER: BIOG-20121 USA					
<b>TELECOMMUNICATION INFORMATION:</b>					
TELEPHONE: 415-433-4150					
TELEFAX: 415-433-8716					
<b>INFORMATION FOR SEQ ID NO: 3:</b>					
<b>SEQUENCE CHARACTERISTICS:</b>					
LENGTH: 886 base pairs					
TYPE: nucleic acid					
STRANDEDNESS: single					
TOPOLOGY: linear					
MOLECULE TYPE: DNA (genomic)					
HYPOTHETICAL: NO					
ANTI-SENSE: NO					
<b>ORIGINAL SOURCE:</b>					
ORGANISM: Chinese cucumber					
<b>IMMEDIATE SOURCE:</b>					
CLONE: alpha-trichosanthin					
<b>FEATURE:</b>					
NAME/KEY: CDS (B) LOCATION: 8..877					
LOCATION: 8..877					
<b>SEQUENCE DESCRIPTION: SEQ ID NO: 3:</b>					
US-08-484-341-3					
<b>Alignment Scores:</b>					
Pred. No.: 1,06e-166 Length: 886					
Score: 1413.00 Matches: 285					
Percent Similarity: 99.31% Conservative: 2					
Best Local Similarity: 98.62% Mismatches: 2					
Query Match: 98.67% Indels: 0					
DB: 3 Gaps: 0					
<b>US-09-905-247-1 (1-289) x US-08-484-341-3 (1-886)</b>					
QY	1	MetIleArgPheLeuValLeuSerLeuLeuThrLeuPheLeuThrProAla	20		
Db	8	ATGATCAGATTCCTTAGTCTCTCTTTGTCTTAATTCACCCCTTCTCTAACAACTCCTGCT	67		
QY	21	ValGluglyAspValSerPheArgLeuSerGlyAlaThrSerSerSerTyrglyValPhe	40		
Db	68	GTTGGAGGGCGATGTTAGCTTCGGTTTATCAGGTGCAACAAGCAGTTCCTATGCAGTTTC	127		
QY	41	IleSerAsnLeuArgLysAlaLeuProAsnGluArgLysLeuTyrglyAspIleProLeuLeu	60		
Db	128	ATTTCAAATCTGAAAAGACTCTTCCAATGAAGAAGAACTGTACGATATCCCTCTGTTA	187		
QY	61	ArgSerSerLeuProGlySerGlnArgTyraLaLeuHisLeuThrAsnTyraLaasp	80		
Db	188	CGTTCCTCTCTCCAGGTTCTCAACGCTACGCAATGATCCATCTCACAAATTTACGCCGAT	247		
QY	81	GluThrIleSerValAlaIleAspValThrAsnValTyrlleMetGlyTyraAgAlagly	100		
Db	248	GAACACCATTTTCAGTGCCCATAGACGTAACGACGCTCTATATATTGGGATATCGCGCTGGC	307		
QY	101	AspThrSerTyrrPhePheAsnGlyAlaSerAlaThrGluAlaIalalytyrValPheLys	120		
Db	308	GATACATCTCTATTTTTTCAACGAGGCTTCTGCAACAGAGCTGCAAAATATGATTCAA	367		
QY	121	AspAlaMetArgLysValThrLeuProTyrrSerGlyAsnTyrrGluArgLeuGlnThrAla	140		
Db	368	GACGCTATGCGAAAAGTTAGCTTCCATATTCGGCAATTTACGAAAGGCTTCAAACTGCT	427		
QY	141	AlsGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThr	160		
Db	428	CGCGGCAAAATAAGGGAATAATTTCCGCTTGGACTCCAGCTTTGGACAGTGCATACC	487		
QY	161	ThrLeuPheTyrrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSer	180		
Db	488	ACTTTGTTTTTACTACCAACGCCAATCTGCTGCGCTCGCACTTATGTTACTTCATTCAGTCG	547		
QY	181	ThrSerGluAlaAlaArgTyrrLysPheIleGluGlnGlnIleGlyLysArgValAspLys	200		
Db	548	ACGCTCTGAGGCTGCGAGGTATAAATTTATTGAGCAACAATTTGGGAAGCGCGTTGACAAA	607		
QY	201	ThrPheLeuProSerLeuAlaIleIleSerLeuGluAsnSerTrpSerAlaLeuSerLys	220		
Db	608	ACCTTCTTACCAAGTTTAGCAATTTAAGTTTGGAAAAATAGTTGGTCTGCTCTCTCCAG	667		
QY	221	GlnIleGlnIleAlaSerThrAsnAsnGlyGlnPheGluSerProValValleuleAsn	240		
Db	668	CAAAATTCAGATGCGGAGTACTAATATGACAGAGTTTGAACCTCTGTTGTGCTTATAAT	727		
QY	241	AlaGlnAsnGlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnIle	260		
Db	728	GCTCAAAACCAACGAGTCATGATAACCAATGTTGATGCTGGAGTTGTHAACCTCCACATC	787		
QY	261	AlaLeuLeuLeuAsnArgAsnAsnMetAlaAlaMetAspAspValProMetThrGln	280		
Db	788	GGTGTGCTGCTGAATCGAAACAATATGCGCAGCATGATGACGATGTTCTTCTATGACACAG	847		
QY	281	SerPheGlyCysGlySerTyrrAlaLeu	289		
Db	848	ACGTTTGGATGTGGAAGTTATGCTAIT	874		
<b>RESULT 5</b>					
US-08-483-502-3					
Sequence 3, Application US/08483502					
Patent No. 6284492					
<b>GENERAL INFORMATION:</b>					
APPLICANT: Donson, Jon					
APPLICANT: Dawson, William O.					
APPLICANT: Grantham, George L.					
APPLICANT: Turpen, Thomas H.					
APPLICANT: Turpen, Ann Myers					
APPLICANT: Garger, Stephen J.					
APPLICANT: Grill, Laurence K.					
<b>TITLE OF INVENTION: RECOMBINANT PLANT VIRAL NUCLEIC ACIDS</b>					
<b>NUMBER OF SEQUENCES: 11</b>					
<b>CORRESPONDENCE ADDRESS:</b>					
ADDRESSEE: Venable, Baetjer, Howard & Civiletti					
STREET: 1201 New York Avenue N.W., Suite 1000					
CITY: Washington					
STATE: DC					
COUNTRY: USA					
ZIP: 20005					
<b>COMPUTER READABLE FORM:</b>					
MEDIUM TYPE: Floppy disk					
COMPUTER: IBM PC compatible					
<b>OPERATING SYSTEM: PC-DOS/MS-DOS</b>					
SOFTWARE: PatentIn Release #1.0, Version #1.25					
<b>CURRENT APPLICATION DATA:</b>					

RESULT 6  
US-09-726-651A-3  
; Sequence 3, Application US/09726651A  
; Patent No. 6448046  
; GENERAL INFORMATION:  
; APPLICANT: Donson, Jon  
; APPLICANT: DAWSON, William O.  
; APPLICANT: GRANTHAM, GEORGE L.  
; APPLICANT: TURPEN, Thomas H.  
; APPLICANT: TURPEN, Ann M.  
; APPLICANT: GARGER, Stephen J.  
; APPLICANT: GRILL, Laurence K.  
; TITLE OF INVENTION: RECOMBINANT ANIMAL VIRUS NUCLEIC ACIDS  
; FILE REFERENCE: 008010023CNS01  
; CURRENT APPLICATION NUMBER: US/09/726, 651A  
; CURRENT FILING DATE: 2002-05-02  
; PRIOR APPLICATION NUMBER: 08/483,502  
; PRIOR FILING DATE: 1995-06-07  
; PRIOR APPLICATION NUMBER: 08/184,237  
; PRIOR FILING DATE: 1994-01-19  
; PRIOR APPLICATION NUMBER: 07/923,692  
; PRIOR FILING DATE: 1992-07-31

```

APPLICATION NUMBER:  US/08/483,502
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:  07/739,143
FILING DATE:
APPLICATION NUMBER:  US 600,244
FILING DATE:  22-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER:  US 641,617
FILING DATE:  16-JAN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER:  US 310,881
FILING DATE:  17-FEB-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER:  US 160,766
FILING DATE:  26-FEB-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER:  US 160,771
FILING DATE:  26-FEB-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER:  US 347,637
FILING DATE:  05-MAY-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER:  US 363,138
FILING DATE:  08-JUN-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER:  US 219,279
FILING DATE:  15-JUL-1988
ATTORNEY/AGENT INFORMATION:
NAME:  Innen, Jeffrey L.
REGISTRATION NUMBER:  28,957
REFERENCE/DOCKET NUMBER:  18604-090574
TELECOMMUNICATION INFORMATION:
TELEPHONE:  202-962-4810
TELEFAX:  202-962-8300
INFORMATION FOR SEQ ID NO:  3:
SEQUENCE CHARACTERISTICS:
LENGTH:  886 base pairs
TYPE:  nucleic acid
STRANDEDNESS:  single
TOPOLOGY:  linear
MOLECULE TYPE:  DNA (genomic)
HYPOTHETICAL:  NO
ANTI-SENSE:  NO
ORIGINAL SOURCE:
ORGANISM:  Chinese cucumber
IMMEDIATE SOURCE:
CLONE:  alpha-trichosanthin
FEATURE:
NAME/KEY:  CDS
LOCATION:  8..877
US-08-483-502-3

Alignment Scores:
Pred. No.:  1,06e-166
Score:  1413.00
Percent Similarity:  99.31%
Best Local Similarity:  98.62%
Query Match:  98.67%
DB:  3

US-09-905-247-1 (1-289) x US-08-483-502-3 (1-886)

QY 1 MetIleA r g p h e L e u v a l l e u s e r l e u l l e L e u t h r L e u P h e L e u T h r P r o A l a 20
Db 8 A T G A T C A G A T T C T T A G T C C T C T T T G C T A A T T C A C C C T C T T C C T A C A A C T C T G C T 67
QY 21 V a l G l u G l y A s p V a l s e r P h e a r g L e u s e r G l y A l a T h r S e r S e r T y r G l y V a l P h e 40
Db 68 G T G G A G G C G A G T T A G C T T C C G T T A T C A G G T G C A A C A A G C A G T T C C T A T G G A G T T T C 127
QY 41 l l e S e r A s n L e u A r g L y s a l a L e u P r o A n G l u A r g L y s L e u T y r A s p l e P r o L e u L e u 60

```

```

; PRIOR APPLICATION NUMBER: 07/600,244
; PRIOR FILING DATE: 1990-10-22
; PRIOR APPLICATION NUMBER: 07/641,617
; PRIOR FILING DATE: 1991-01-16
; PRIOR APPLICATION NUMBER: 07/737,899
; PRIOR FILING DATE: 1991-07-26
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 886
; TYPE: DNA
; ORGANISM: Chinese Cucumber
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (8)...(877)
; OTHER INFORMATION: Clone:alpha-trichosanthin
US-09-726-651A-3

Alignment Scores:
Pred. No.: 1,06e-166 Length: 886
Score: 1413.00 Matches: 285
Percent Similarity: 99.31% Conservatave: 2
Best Local Similarity: 98.62% Mismatches: 2
Query Match: 98.67% Indels: 0
DB: 4 Gaps: 0

US-09-905-247-1 (1-289) x US-09-726-651A-3 (1-886)

QY 1 MetIleAtrGpheLeuValLeuSerLeuLeuLeuThrLeuPheLeuThrThrProAla 20
DB 8 ATGATCAGATTCTTAGTCCTCTCTTTGGCTAAATTCACCCCTCTTCCTTAAACAACCTCGCT 67
QY 21 ValGluGlyAspValSerPheArgLeuSerGlyAlaThrSerSerSerTyrGlyValPhe 40
DB 68 GTGGAGGGCGATGTTAGCTTCGGTTATCAGGTGCACCAAGCAGTTCTTCATGGAGTTTC 127
QY 41 IleSerAsnLeuArgGlyAlaLeuProAsnGluArgLysLeuTyrAspIleProLeuLeu 60
DB 128 ATTTCAAATCTGAGAAAAGCTCTCCAAATGAAAGGAAACTGTACGATATCCCTCTGTTA 187
QY 61 ArgSerSerLeuProGlySerGlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAsp 80
DB 188 CGTTCCTCTCTCCAGGTCTCTCAACGCTACGCAITTGATCCATCTCACAAATATACGCCGAT 247
QY 81 GluThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGly 100
DB 248 GAAACCAATTTCAAGTGGCCATAGACGTACGAACGCTATATTTATGGGATATCGCGTGGC 307
QY 101 AspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPheLys 120
DB 308 GATACATCTCTATTTTTTCAACGAGGCTTCTGCAACAGAAAGCTGCAAAATATGTAATCAA 367
QY 121 AspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAla 140
DB 368 GACGCTATGCGAAAAGTTACGCTTCCATATTTCTGGCAATTCGAAAGGCTTCAAACTGCT 427
QY 141 AlsGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThr 160
DB 428 GCGGGCAAAATAAGGGAAATATTCGGCTTGGACTCCCGAGCTTTGGACAGTGCCATTACC 487
QY 161 ThrLeuPheTyrTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSer 180
DB 488 ACTTTGTTTACTACAAACGCCAAATTCGTCTCGTCGGCACATTATGGTACTCATTCAGTCG 547
QY 181 ThrSerGluAlaAlaArgTyrLysPheIleGluGlnGlnIleGlyLysArgValAspLys 200
DB 548 ACGTCTGAGGCTCGGAGGTATAAATTTATTGAGCAACAAATTTGGAGACGGGTTGACAAA 607
QY 201 ThrPheLeuProSerLeuAlaIleIleSerLeuGluAsnSerTrpSerAlaLeuSerLys 220
DB 608 ACCTTCCTCAAGTTTAGCAATATTAAGTTTGGAAAATAGTTCGTCTCTCTCCAAAG 667
QY 221 GlnIleGlnIleAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuIleAsn 240

```

```

QY 21 ValGluGlyAspValSerPheArgLeuSerGlyAlaThrSerSerTyrGlyValPhe 40
Db 104 GTTCAGGGCGATGTTAGCTTCGGTTATTCAGGGTGTACACCACTCTTCAAGAGTTTC 163
QY 41 IleSerAsnLeuArgLysAlaLeuProAsnGluArgLysLeuTyrAspIleProLeuLeu 60
Db 164 ATTAAAAATCTGAGAGAGCTCTTCCTACAGAAAGAAAGTGTACATATACCGCTATTA 223
QY 61 ArgSerSerLeuProGlySerGlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAsp 80
Db 224 CGTTCAGTATTTTCAGGTTTCAGGAGCGCTACACATTTCTCCATCTTCACAAATTCAGCGGAT 283
QY 81 GluThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGly 100
Db 284 GAAACCATCTCAGTGGCAGTAGCGTAAACAACTCTATATTTATGGGTATCTTGGCGGT 343
QY 101 AspThrSerTyrPhePheAsnGlyAlaThrGluAlaAlaLysTyrValPheLys 120
Db 344 GATGTGCTCTATTTTTCAGGAGCGCTTCAGCAACAGAGCTGCACAAATTCGTATTCAAA 403
QY 121 AspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAla 140
Db 404 GATGCTTAAGAAAAAGTGGAGCTTCCATATTCAGGCAATTCAGAAAGCTTCAAACTGCT 463
QY 141 AlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThr 160
Db 464 GCAGGAAAAATAAGAGAAAAATATTCCTGACTTCCAGCTTCCAGCTTGGACAGTCCCATACC 523
QY 161 ThrLeuPheTyrTyrAsnAlaSerAlaAlaSerAlaLeuMetValLeuIleGlnSer 180
Db 524 ACTTTGTTATTTACTACCGCCAGTTCTCGGGCTTCTGCACTTCTTGTTACTTCAATCC 583
QY 181 ThrSerGluAlaAlaArgTyrIlePheLeuGlnGlnIleGlyIleArgValAspLys 200
Db 584 ACGGCTGAATCTGCAAGGTATATAATTTATTAACCAACAAATTTGGAAGCGTGTAGACAAA 643
QY 201 ThrPheLeuProSerLeuAlaIleIleSerLeuGluAsnSerTyrPheAlaLeuSerLys 220
Db 644 ACTTTTTCAGGAGTTTACCACTATTTAGTTTGGAAATAATTTGCTGCTCTGCTCCAG 703
QY 221 GlnIleGlnIleAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuIleAsn 240
Db 704 CAATTCAGATGAGCAGTACCAATAATGACAAATTTGAGAGTCTCTGCTGCTTATAGAT 763
QY 241 AlaGlnAsnGlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnIle 260
Db 764 GGTAAACAACCAACGAGTCTCTATAACCAATGCTAGTGTGCTGAGTGTGTAACCTCCACATA 823
QY 261 AlaLeuLeuLeuAsnArgAsnAsnMetAlaAlaMetAspAspValProMetThrGln 280
Db 824 GGGTTCGCTGCTAAACAGAAATAATTTGGCCATTTGGAGAGGACATTTCTATGACACTC 883
QY 281 ---SerPheGlyCysGlySerTyrAlaLeu 289
Db 884 ATCGGCTTTCAGACATGCACTTTATGTATTA 913

```

RESULT 8

```

US-08-597-731-1
; Sequence 1, Application US/08597731
; Patent No. 5932447
; GENERAL INFORMATION:
; APPLICANT: Siegal, Clay B.
; TITLE OF INVENTION: Cloning and Expression of a Gene
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bristol-Myers Squibb Company
; STREET: 3005 First Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:

```

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/597,731
; FILING DATE: 07-FEB-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/245,754
; FILING DATE: 06-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Poor, Brian W.
; REGISTRATION NUMBER: 32,928
; REFERENCE/DOCKET NUMBER: ON0124
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 728-4800
; TELEFAX: (206) 727-3601
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1094 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Bryonia dioica
; US-08-597-731-1
; Alignment Scores:
; Pred. No.: 1.55e-135 Length: 1094
; Score: 1164.50 Matches: 235
; Percent Similarity: 90.00% Conservative: 26
; Best Local Similarity: 81.03% Mismatches: 28
; Query Match: 81.32% Indels: 1
; Gaps: 2
; US-09-905-247-1 (1-289) x US-08-597-731-1 (1-1094)
QY 1 MetIleArgPheLeuValLeuSerLeuLeuLeuLeuLeuThrLeuPheLeuThrProAla 20
Db 44 ATGATCAAAATGTTAGTCTTTGGTTGTTGTTATTTCTCACCATATTTCTCAAAATCTCAACT 103
QY 21 ValGluGlyAspValSerPheArgLeuSerGlyAlaThrSerSerSerTyrGlyValPhe 40
Db 104 GTTCAGGGCGATGTTAGCTTCGGTTATTCAGGTTGTACAAACCACTCTTATGGAGTTTC 163
QY 41 IleSerAsnLeuArgLysAlaLeuProAsnGluArgLysLeuTyrAspIleProLeuLeu 60
Db 164 ATTAAAAATCTGAGAGAGCTCTTCATACGAAAGAAAGTGTACATATACCGCTATTA 223
QY 61 ArgSerSerLeuProGlySerGlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAsp 80
Db 224 CGTTCAGTATTTTCAGGTTTCAGGAGCGCTACACATTTCTCCATCTTCACAAATTCAGCGGAT 283
QY 81 GluThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGly 100
Db 284 GAAACCATCTCAGTGGCAGTAGCGTAAACAACTCTATATTTATGGGTATCTTGGCGGT 343
QY 101 AspThrSerTyrPhePheAsnGlyAlaThrGluAlaAlaLysTyrValPheLys 120
Db 344 GATGTGCTCTATTTTTCAGGAGCGCTTCAGCAACAGAGCTGCACAAATTCGTATTCAAA 403
QY 121 AspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAla 140
Db 404 GATGCTTAAGAAAAAGTGGAGCTTCCATATTCAGGCAATTCAGAAAGCTTCAAACTGCT 463
QY 141 AlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThr 160
Db 464 GCAGGAAAAATAAGAGAAAAATATTCCTGACTTCCAGCTTGGACAGTCCCATACC 523
QY 161 ThrLeuPheTyrTyrAsnAlaSerAlaAlaSerAlaLeuMetValLeuIleGlnSer 180

```

```

Db      524  ACTTTGTTTACTACACCGCAGTCTGCGCTTCTGCACTTCTGTACTCAATCAATCC 583
QY      181  ThrSerGluAlaAlaArgTyrLysPheIleGlnGlnIleGlyLysArgValAspLys 200
Db      584  ACGGCTGAATCTCAAGATATAATTTATTGAAACAAATTTGAAAGCGGTAGACAAA 643
QY      201  ThrPheLeuProSerLeuAlaIleSerLeuGluAsnSerTrpSerAlaLeuSerLys 220
Db      644  ACTTTTTCACCAAGTTTAGCACTATTAGTTTGAAATAATTTGGTCTGCTGCTCAAG 703
QY      221  GlnIleGlnIleAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuIleAsn 240
Db      704  CAATTCAGATACCCAGTACCAATAATGGAACAATTTGAGAGTCTGTTGTCTTAGAT 763
QY      241  AlaGlnAsnGlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnIle 260
Db      764  GGTAACAACCAACGAGTCTCTATAACCAATGCTAGTCTGCTGAGTTGTAACCTCCACATA 823
QY      261  AlaLeuLeuLeuAsnArgAsnAsnMetAlaAlaMetAspAspValProMetThrGln 280
Db      824  GCGTGTCTGCTTAACAGAAATAATATTGACAGCAATTTGGAGAGACATTTCTATGACATC 883
QY      281  ---SerPheGlyCysGlySerTyrAlaLeu 289
Db      884  ATCGGCTTTGAACATGGCACTTTATGGTATA 913

```

## RESULT 9

```

US-08-902-486-14
; Sequence 14, Application US/08902486
; Patent No. 6140075
; GENERAL INFORMATION:
; APPLICANT: Russel, David R.
; TITLE OF INVENTION: METHOD FOR PRODUCING ANTIBODIES AND
; TITLE OF INVENTION: PROTEIN TOXINS IN PLANT CELLS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE:
; CLASSIFICATION: US/08/902,486
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 670513.90261
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1499 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..1493
US-08-902-486-14
Alignment Scores:

```

```

Pred. No.: 5,08e-124 Length: 1499
Score: 1074.00 Matches: 217
Percent Similarity: 90.46% Conservative: 20
Best Local Similarity: 82.82% .Mismatches: 21
Query Match: 75.00% Indels: 4
DB: 3 Gaps: 1

US-09-905-247-1 (1-289) x US-08-902-486-14 (1-1499)
QY      24  AspValSerPheArgLeuSerGlyAlaThrSerSerSerTyrGlyValPheIleSerAsn 43
Db      6  GATGTGAGCTTTCGTTTATCAGGTGCTACAAACCATCTCTATGAGATTTTCATTAATAAT 65
QY      44  LeuArgGlyAlaLeuProAsnGluArgLysLeuTyrAspIleProLeuLeuArgSerSer 63
Db      66  CTGAGAGAGCTCTCCATACGAAAGGAAAGTGTACAATATACCGCTATTACGTTCAAGT 125
QY      64  LeuProGlySerGlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAspGluThrIle 83
Db      126  AITTCAGGTTTCAGGACGCTACACATTCTCCATCTCACAAATTACGCGGATGAACCATC 185
QY      84  SerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGlyAspThrSer 103
Db      186  TCAGTGGCAGTAGACGTAAACAAACGCTATATTATCGGGTATCTTCGCGGTATGTGTCC 245
QY      104  TyrPhePheAsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPheLysAspAlaMet 123
Db      246  TATTTTTCACAGAGCTTCAGCAACAGAGCTGCAAAATTCGTATTCAAGATGCTAAG 305
QY      124  ArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAlaAlaGlyLys 143
Db      306  AAAAAGTACGCTTCATATTTCAGGCAATTACGAAAGCTTCAAACTGCTCAGGAAA 365
QY      144  IleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThrThrLeuPhe 163
Db      366  ATAAGAGAAAATATTCCACTTGGACTCCAGCTTTGGACAGTGGCATTTACCACTTTGTAT 425
QY      164  TyrTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSerThrSerGlu 183
Db      426  TACTACACCGCCAGTCTCGCGCTTCTGCACTTCTTGTAATCTTCAATCCACGCGTGAA 485
QY      184  AlaAlaArgTyrLysPheIleGlnGlnIleGlyLysArgValAspLysThrPheLeu 203
Db      486  TCTGCAAGGTATAAATTTATTGAACAATAATGAAAGCGGTGTAGACAAAACCTTTTFA 545
QY      204  ProSerLeuAlaIleIleSerLeuGluAsnSerTrpSerAlaLeuSerLysGlnIleGln 223
Db      546  CCAAGTTTAGCAACTATTAGTTTGGAAAATAATTTGGTCTGCTCTGCTGCCAACCAATTGAG 605
QY      224  IleAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuIleAsnAlaGlnAsn 243
Db      606  ATAGCCAGTACCAATAATGGAACAATTTGAGAGTCTCTGTTGTGCTTATAGATGGAACAAC 665
QY      244  GlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnIleAlaLeuLeu 263
Db      666  CAACGAGTCTCTATAACCAATGCTAGTGTGAGTTGTAACCTCCAAATAGCGTGTGCTG 725
QY      264  LeuAsnArgAsnAsnMetAlaAlaMetAspAsp-----AspValProMetThr 279
Db      726  CTAACAGAAAATAATATTGACGCAATGATGCTACCAAGGCCATGGATGCTGTGTATGACC 785
QY      280  GlnSer 281
Db      786  CAAAAT 791

RESULT 10
US-08-902-486-6
; Sequence 6, Application US/08902486
; Patent No. 6140075
; GENERAL INFORMATION:
; APPLICANT: Russel, David R.
; APPLICANT: Fuller, James T.
; TITLE OF INVENTION: METHOD FOR PRODUCING ANTIBODIES AND

```

Wed Sep 15 10:32:02 2004

TITLE OF INVENTION: PROTEIN TOXINS IN PLANT CELLS

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESSEE: Quarles & Brady

STREET: 1 South Pinckney Street

CITY: Madison

STATE: WI

COUNTRY: US

ZIP: 53701-2113

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/902,486

FILING DATE:

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: Seay, Nicholas J.

REGISTRATION NUMBER: 27386

REFERENCE/DOCKET NUMBER: 670513.90261

TELECOMMUNICATION INFORMATION:

TELEPHONE: 608-251-5000

TELEFAX: 608-251-9166

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 758 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 3..748

US-08-902-486-6

Alignment Scores:

Pred. No.: 2,99e-123

Score: 1064.00

Percent Similarity: 93.52%

Best Local Similarity: 85.83%

Query Match: 74.30%

DB: 3

US-09-905-247-1 (1-289) x US-08-902-486-6 (1-758)

QY 24 AspValSerPheArgLeuSerGlyAlaThrSerSerSerTyrGlyValPheLeSerAsn 43  
 Db 6 GATGTGAGCTTTCGTTTATCAGTGCTACACCATCTCTGAGTTCATTAATAAT 65  
 QY 44 LeuArgGlyAlaLeuProAsnGluA:GlyLeuTyrAspIleProLeuLeuArgSerSer 63  
 Db 66 CTGAGAGAGCTCTCCATACGAAAGGAGTGTACATATATACCGCTATTACGTTCAAGT 125  
 QY 64 LeuProGlySerGlnArgTyrAlaLeuLeuHisLeuThrAsnTyrAlaAspGluThrIle 83  
 Db 126 ATTTCAGGTTCAGGAGCTACACATTACTCTCATCTCACAATTCGCGGATGAACCATC 185  
 QY 84 SerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGlyAspThrSer 103  
 Db 186 TCAGTGCGAGTACAGCTAACACGCTATATTATGGGTATCTTCCGCGGTATGTCTCC 245  
 QY 104 TyrPhePheAsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPheLysAspAlaMet 123  
 Db 246 TATTTTTCACGAGGCTTCAGCAACAGAGCTGCAAAATTCGTATTCAAGATGCTAAG 305  
 QY 124 ArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAlaAlaGlyLys 143  
 Db 306 AAAAAAGTACGCTTCCATATTCAGGCAATTCAGAAAGGCTTCAACTGCTGCAGGAAA 365  
 QY 144 IleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThrThrLeuPhe 163

Db 366 ATAAGAGAAAATATTCACCTTGACTCCAGCTTTGGACAGTCCATTACCACCTTTGTAT 425  
 QY 164 TLTyTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSerThrSerGlu 183  
 Db 426 TACTACACGCCAGTTCGCGGCTTCGCACTTCTTGACTCATTCATCAATCCACGGCTGAA 485  
 QY 184 AlaAlaArgTyrLysPheIleGluGlnGlnIleGlyLysArgValAspLysThrPheLeu 203  
 Db 486 TCTGCAAGGTATAAATTTATTGAACACACAAATTTGGAAGCGGTGAGACAAACATTTTITA 545  
 QY 204 ProSerLeuAlaIleIleSerLeuGluAsnSerTipSerAlaLeuSerLysGlnIleGln 223  
 Db 546 CCAAGTTTACCACTATTAGTTTGGAAAAATAATTTGGTCTGCTCTGCTCCAGCAATTCAG 605  
 QY 224 IleAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuIleAsnAlaGlnAsn 243  
 Db 606 ATAGCCAGTACCAATAATGACAAATTTGAGAGTCTCTGTGTGTATATAGATGGTAACAAC 665  
 QY 244 GlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnIleAlaLeuLeu 263  
 Db 666 CAACGAGTCTCTATAACCAATGTAGTCTCGAGTGTGTAGTCTCGAGTGTGTAGTCTGCTG 725  
 QY 264 LeuAsnArgAsnAsnMetAla 270  
 Db 726 CTAACAGAAATAATATGCA 746  
 RESULT 11  
 US-07-901-707-13  
 ; Sequence 13, Application US/07901707  
 ; Patent No. 5376546  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bernhard, Susan L.  
 ; APPLICANT: Better, Marc D.  
 ; APPLICANT: Carroll, Steve F.  
 ; APPLICANT: Lane, Julie A.  
 ; TITLE OF INVENTION: Materials Comprising and Methods of  
 ; USE OF INVENTION: Composition and Use for Ribosome-Inactivating Proteins  
 ; NUMBER OF SEQUENCES: 57  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
 ; ADDRESSEE: Bicknell  
 ; STREET: Two First National Plaza, 20 South Clark  
 ; STREET: Street  
 ; CITY: Chicago  
 ; STATE: Illinois  
 ; COUNTRY: USA  
 ; ZIP: 60603  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/07/901,707  
 ; FILING DATE: 19920619  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/787,567  
 ; FILING DATE: 04-NOV-1991  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: No. 5376546and, Greta E.  
 ; REGISTRATION NUMBER: 35,302  
 ; REFERENCE/DOCKET NUMBER: 27129/30910  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (312) 346-5750  
 ; TELEFAX: (312) 984-5750  
 ; TELEX: 25-3856  
 ; INFORMATION FOR SEQ ID NO: 13:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 913 base pairs  
 ; TYPE: NUCLEIC ACID  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear

```

; GENERAL INFORMATION:
; APPLICANT: Bernhard, Susan L.
; APPLICANT: Better, Marc D.
; APPLICANT: Carroll, Stephen F.
; APPLICANT: Lane, Julie A.
; APPLICANT: Lei, Shau-ping
; TITLE OF INVENTION: Materials Comprising and Methods of
; TITLE OF INVENTION: Preparation and Use for Ribosome-Inactivating Proteins
; NUMBER OF SEQUENCES: 101
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Bicknell
; STREET: Two First National Plaza, 20 South Clark
; STREET: Street
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/988,430
; FILING DATE: 19921209
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/901,707
; FILING DATE: 19-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/787,567
; FILING DATE: 04-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5416202and, Greta E.
; REGISTRATION NUMBER: 35302
; REFERENCE/DOCKET NUMBER: 31133
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 346-5750
; TELEFAX: (312) 984-9740
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 913 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-07-988-430-13

Alignment Scores:
Pred. No.: 3,5e-81 Length: 913
Score: 728.00 Matches: 156
Percent Similarity: 72.83% Conservative: 37
Best Local Similarity: 58.87% Mismatches: 70
Query Match: 50.84% Indels: 3
DB: 1 Gaps: 2

US-09-905-247-1 (1-289) x US-07-988-430-13 (1-913)

QY 1 MetIleArgPheLeuValLeuSerLeuLeuLeuLeuThrLeuPheLeuThrProAla 20
Db 10 ATCGTGAATGCTTACTACTTCTCTTTTAAATTCGCCATCTTCATTGGTGCTCTACT 69
QY 21 ValGluGlyAspValSerPheArgLeuSerGlyAlaThrSerSerSerTyrGlyValPhe 40
Db 70 GCCAAGGCGATGTTAACTTCGATTTTCGACTGCCACTGCAAAACCTACACAAATTT 129
QY 41 IleSerAsnLeuArgLysAlaLeuProAsnGluArgLysLeuTyrAspIleProLeuLeu 60
Db 130 ATCGAAGATTTCAGGCGACTCTTCATTAGCCATAAGTGTAATGATATACCTCTACTG 189
QY 61 ArgSerSerLeuProGlySerGlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAsp 80

```

```

/ APPLICATION NUMBER:  US 07/901,707
/ FILING DATE: 19-JUN-1992
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER:  US 07/787,567
/ FILING DATE: 04-NOV-1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Meyers, Thomas C.
/ REGISTRATION NUMBER: P-36,989
/ REFERENCE/DOCKET NUMBER: 31394
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 312/474-6300
/ TELEFAX: 312/474-0448
/ TELEX: 25-3856
/ INFORMATION FOR SEQ ID NO: 13:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 913 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: cdna
/ US-08-425-336-13

Alignment Scores:
Pred. No.: 3 5e-81 Length: 913
Score: 728.00 Matches: 156
Percent Similarity: 72.83% Conservatives: 37
Best Local Similarity: 58.87% Mismatches: 70
Query Match: 50.84% Indels: 3
DB: 1 Gaps: 2

US-09-905-247-1 (1-289) x US-08-425-336-13 (1-913)
QY 1 MetIleArgPheLeuValLeuSerLeuLeuLeuThrLeuPheLeuThrThrProAla 20
DB 10 ATCGTGAATGCITACTACTTCITCTTTTAAATTATCGCCATCTTCATTTGGTGTCTACT 69
QY 21 ValGluGlyAspValSerPheArgLeuSerGlyAlaThrSerSerSerTyrGlyValPhe 40
DB 70 GCCAAAGGCGATGTTTACTTTCGATTTCTCGACTGCCACTGCCAAACCTCACAAAATTT 129
QY 41 IleSerAsnLeuArgLysAlaLeuProAsnGluArgLysLeuTyrAspIleProLeuLeu 60
DB 130 ATCGAAGATTTCAGGGCGACTTTCCTCAATTTAGCCATAAAGTGATGATATACCTCTACTG 189
QY 61 ArgSerSerLeuProGlySerGlnArgTyrAlaLeuLeuHisLeuThrAsnTyrAlaAsp 80
DB 190 TATTCACATATTCGACTCCGACTCCAGACGTTTCATCTCTCGATCTTACAAAGTTATGCATAT 249
QY 81 GluThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGly 100
DB 250 GAAACCATCTCGGTGGCGCATAGATGTGACGAACGTTTATGTGTGGCGTATCGCACCCGCG 309
QY 101 AspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPheLys 120
DB 310 GATGTATCTCTACTTTTTTAAA---GAATCTCTCTCTGAAGCTTATAACATCTCTATTCAA 366
QY 121 AspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAla 140
DB 367 ---GGTACGCGGAAAATTTACTGCCATATACCGGTAAATATGAAAACTCTTCAAACTGCT 423
QY 141 AlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThr 160
DB 424 GCACAAAAAAGAGAGAAATATTGATCTTGACTCCCTGCCCTTCGAGTAGTCCCAATTACC 483
QY 161 ThrLeuPheTyrTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSer 180
DB 484 ACATTGTTTATTACAAATGCCCAATCTGCTCTCTTCGTGATTCGTTGTACTAATCCAGACG 543
QY 181 ThrSerGluAlaAlaArgTyrIlyPheIleGluGlnGlnIleGlyLysArgValAspLys 200
DB 544 ACTGCAGAGAGCTGCAGATTAAAGTATATCGAGCGACACGCTGTGCTAAAGTATGTTGCCACT 603
QY 201 ThrPheLeuProSerLeuAlaIleIleSerLeuGluAsnSerTrpSerAlaLeuSerLys 220

```

RESULT 13  
 US-08-425-336-13  
 ; Sequence 13, Application US/08425336  
 ; Patent No. 5621083  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Better, Marc D.  
 ; APPLICANT: Carroll, Stephen F.  
 ; APPLICANT: Studnika, Gary M.  
 ; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
 ; TITLE OF INVENTION: Proteins  
 ; NUMBER OF SEQUENCES: 140  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
 ; STREET: 6300 Sears Tower, 233 South Wacker Drive  
 ; CITY: Chicago  
 ; STATE: Illinois  
 ; COUNTRY: USA  
 ; ZIP: 60606-6402  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/425,336  
 ; FILING DATE: 18-APR-1995  
 ; CLASSIFICATION: 530  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/064,691  
 ; FILING DATE: 12-MAY-1993



604 AACCTTAAAGCAATCTAGCCATCATAGCTTGGAAATCAATGGTCTGCTCTCCAA- 662  
221 GlnIleGlnIleAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuIleAsn 240  
663 CAATCTTTTGGCGCAGAAATCAAGGAGGAAATTTAGAAATCTGTCGACCTTATAAA 722  
241 AlaGlnAsnGlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnIle 260  
723 CCTACCGGGGAGCGTTTCAAGTAACCAATGTTGATTTCAGATGTTGAAAAAGGTAATATC 782  
261 AlaLeuLeuLeuAsn 265  
783 AAATCTCTGCTGAAC 797

RESULT 14

US-08-488-113B-13  
; Sequence 13, Application US/08488113B  
; Patent No. 5744580  
; GENERAL INFORMATION:  
; APPLICANT: Better, Marc D.  
; APPLICANT: Carroll, Stephen F.  
; APPLICANT: Studnika, Gary M.  
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
; NUMBER OF SEQUENCES: 169  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
; STREET: 500 West Madison Street, 34th floor  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60661  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/488,113B  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/425,336  
; FILING DATE: 18-APR-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/064,691  
; FILING DATE: 12-MAY-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/988,430  
; FILING DATE: 09-DEC-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/901,707  
; FILING DATE: 19-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/787,567  
; FILING DATE: 04-NOV-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McNicholas, Janet M.  
; REGISTRATION NUMBER: 32,918  
; REFERENCE/DOCKET NUMBER: 11022US07/200-70.P3.C2A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/707-8889  
; TELEFAX: 312/707-9155  
; TELETYPE: 650 388-1248  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 913 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
US-08-488-113B-13

Alignment Scores:  
Pred. No.: 3.5e-81 Length: 913  
Score: 728.00 Matches: 156  
Percent Similarity: 72.83% Conservative: 37  
Best Local Similarity: 58.87% Mismatches: 70  
Query Match: 50.84% Indels: 3  
DB: 1 Gaps: 2

US-09-905-247-1 (1-289) x US-08-488-113B-13 (1-913)

QY 1 MetIleArgPheLeuValLeuSerLeuLeuIleLeuThrLeuPheLeuThrProAla 20  
Db 10 ATGGTGAATGCTTACTACTTCTTTTAAATATGCCATCTCAATGGTCTGCTCTACT 69  
QY 21 ValGluGlyAspValSerPheArgLeuSerGlyAlaThrSerSerSerGlyValPhe 40  
Db 70 GCCAAAGGCGATGTTAACTTCGATTTGTCGACTGCCACTGCAAAACCTACACAAATTT 129  
QY 41 IleSerAsnLeuArgLysAlaLeuProAsnGluArgLysLeuThrAspIleProLeu 60  
Db 130 ATCGAAGATTTTCAGGCGACTCTTCCATTTAGCCATAAAAGTGTATGATATACCTCTACTG 189  
QY 61 ArgSerSerLeuProGlySerGlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAsp 80  
Db 190 TATTCACATATTTCCGACTCCAGACGTTTCATACCTCTCGATCTTACAAAGTTATGCAAT 249  
QY 81 GluThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGly 100  
Db 250 GAAACCATCTCGTGGCCATAGATGTGACGAACGTTTATGTTGTGGCGTATGCGACCCGC 309  
QY 101 AspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPheLys 120  
Db 310 GATGATCTCTACTTTTAAA---GAATCTCTCTCTGAAGCTTATAACATCTCTTCAA 366  
QY 121 AspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAla 140  
Db 367 ---GGTACGCGGAAATTTACATGCTGCTATACCGGTAATTTAGAAAATCTTCAAATGCT 423  
QY 141 AlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThr 160  
Db 424 GCACACAAATAAGAGAGATATTGATCTTGGACTCCCTGCTTGGTAGTGGCCATTACC 483  
QY 161 ThrLeuPheTyrTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSer 180  
Db 484 ACATTTGTTTATTAACAATGCCCAATCTGCTCTCTTGCATTTGCTGTACTAATCCAGACG 543  
QY 181 ThrSerGluAlaAlaArgTyrLysPheIleGluGlnGlnIleGlyLysArgValAspLys 200  
Db 544 ACTGAGAGCTGCAAGATTTAAGTATATACGAGGACACAGTTGTAGTATGTTGCCACT 603  
QY 201 ThrPheLeuProSerLeuAlaIleIleSerLeuGluAsnSerTrpSerAlaLeuSerLys 220  
Db 604 AACTTTAAGCCAAATCTAGCCATCATAGCTTTGAAATCAATGCTGCTCTCTCCAA- 662  
QY 221 GlnIleGlnIleAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuIleAsn 240  
Db 663 CAATCTTTTGGCGCAGAAATCAAGGAGGAAATTTAGAAATCTGTCGACCTTATAAA 722  
QY 241 AlaGlnAsnGlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnIle 260  
Db 723 CCTACCGGGGAGCGTTTCAAGTAACCAATGTTGATTTCAGATGTTGAAAAAGGTAATATC 782  
QY 261 AlaLeuLeuLeuAsn 265  
Db 783 AAATCTCTGCTGAAC 797

RESULT 15

US-08-477-484B-13  
; Sequence 13, Application US/08477484B  
; Patent No. 5756699  
; GENERAL INFORMATION:  
; APPLICANT: Better, Marc D.

Wed Sep 15 10:32:02 2004

Db 130 ATCGAAGATTTCAGGCGGACTCTTCATTTAGCCATAAAGTGATGATATACCTCTACTG 189  
 Qy 61 ArgSerSerLeuProGlySerGlnArgTyrAlaLeuLeuHisLeuThrAsnTyrAlaAsp 80  
 Db 190 TATTCCACTATTTCGAGCTCCAGACGTTTCATACCTCTCGATCTTCAAGTTTATGATAT 249  
 Qy 81 GluThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGly 100  
 Db 250 GAAACCATCTCGGTGGCCATAGATGACGAAAGCTTTATGTTGGGTATCGCACCCTGC 309  
 Qy 101 AspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPheLys 120  
 Db 310 GATGTATCTTACTTTTAAAA---GAATCTCTCTCTGAGCTTATAACATCTTATTCAAA 366  
 Qy 121 AspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAla 140  
 Db 367 ---GGTACGCGGAAATTCACATGCCATATACCGGTAAATTATGAAATCTTCAAACTGCT 423  
 Qy 141 AlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThr 160  
 Db 424 GCACACAAAATAAGAGAGAATATTGATCTTGAGACTCCCTGCTTGGTGTAGTAGCCATTACC 483  
 Qy 161 ThrLeuPheTyrTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSer 180  
 Db 484 ACATTGTTTATTAACATGCCAATCTGCTCTCTGCTTGTACTTAAATCCAGACG 543  
 Qy 181 ThrSerGluAlaAlaArgTyrLysPheIleGluGlnIleGlyLysArgValAspLys 200  
 Db 544 ACTGAGAAGCTGCAAGATTTAAGTATATCGAGGACACACGTTGCTAAGTATGTTGCCACT 603  
 Qy 201 ThrPheLeuProSerLeuAlaIleSerLeuGluAsnSerTyrPheSerAlaLeuSerLys 220  
 Db 604 AACTTTAAGCCAAATCTAGCCATCATAGCTTGGAAATCAATGGTCTGCTCTCTCCAA- 662  
 Qy 221 GlnIleGlnIleAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuIleAsn 240  
 Db 663 CAATCTTTTGGCGCAGAAATTCAGAAATTTAGAAATCTCTGCTGACCTTATAAAA 722  
 Qy 241 AlaGlnAsnGlnArgValThrIleThrAsnValAlaGlyValValThrSerAsnIle 260  
 Db 723 CCTACCGGGGAAACGGTTTCAAGTAACTAATGTTGATTTCAGATGTTGTAAGGTAATATC 782  
 Qy 261 AlaLeuLeuLeuAsn 265  
 Db 783 AAATCTCTGCTGAAC 797

Search completed: September 14, 2004, 21:58:14  
 Job time : 93 secs

APPLICANT: Carroll, Stephen F.  
 APPLICANT: Studnka, Gary M.  
 TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
 NUMBER OF SEQUENCES: 169  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
 STREET: 500 West Madison Street, 34th floor  
 CITY: Chicago  
 STATE: Illinois  
 COUNTRY: USA  
 ZIP: 60661  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/477,484B  
 FILING DATE: 07-JUN-1995  
 CLASSIFICATION: 530  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/425,336  
 FILING DATE: 18-APR-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/064,691  
 FILING DATE: 12-MAY-1993  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/988,430  
 FILING DATE: 09-DEC-1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/901,707  
 FILING DATE: 19-JUN-1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/787,567  
 FILING DATE: 04-NOV-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: McNicholas, Janet M.  
 REGISTRATION NUMBER: 32,918  
 REFERENCE/DOCKET NUMBER: 11022US07/200-70.P3.C2A  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 312/707-8889  
 TELEFAX: 312/707-9155  
 TELEX: 650 388-1248  
 INFORMATION FOR SEQ ID NO: 13:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 913 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 US-08-477-484B-13

Alignment Scores:  
 Pred. No.: 3 5e-81 Length: 913  
 Score: 728.00 Matches: 156  
 Percent Similarity: 72.83% Conservative: 37  
 Best Local Similarity: 58.87% Mismatches: 70  
 Query Match: 50.84% Indels: 3  
 DB: 1 Gaps: 2

US-09-905-247-1 (1-289) x US-08-477-484B-13 (1-913)

Qy 1 MetIleArgPheLeuValLeuSerLeuLeuLeuThrLeuPheLeuThrProAla 20  
 Db 10 ATGGTGAAGTCTTACTACTCTTTTAAATATATCCCAATCTTATGTTGTTCTTACT 69  
 Qy 21 ValGluGlyAspValSerPheArgLeuSerGlyAlaThrSerSerTyrGlyValPhe 40  
 Db 70 GCCAAGCGGATGTTAACTTCGATTGTCGACTGCCACTGCAAAACCTACACAAATTT 129  
 Qy 41 IleSerAsnLeuArgLysAlaLeuProAsnGluArgLysLeuTyrAspIleProLeuLeu 60

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 14, 2004, 17:15:36 ; Search time 385 Seconds  
(without alignments)  
3188.907 Million cell updates/sec

Title: US-09-905-247-1  
Perfect score: 1432  
Sequence: 1 MIRELVLSLLTLFLTPA.....AAMDVVPMQSGGSGVAL 289

Scoring table:

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-Q/cgn2 1/USPTO.spool\_p/US0905247/runat\_10092004\_103933\_4703/app\_query.fasta\_1.455  
-DB=N Geneseq\_29Jan04 -QWT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum2 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US0905247 @CGN 1 708 @runat\_10092004\_103933\_4703 -NCPU=6 -ICPU=3  
-NO MMAP -LARGESQRY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq\_29Jan04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002s:\*  
7: Geneseqn2003as:\*  
8: Geneseqn2003bs:\*  
9: Geneseqn2003cs:\*  
10: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1416	98.9	870	AAH41473	Aah41473 Trichosan
2	1413	98.7	886	AAQ65573	Aaq65573 Alpha-tri
3	1413	98.7	886	AAT47094	Aat47094 Chinese c
4	1413	98.7	886	AAT61376	Aat61376 Chinese c
5	1413	98.7	886	AAX29143	Aax29143 Chinese c
6	1413	98.7	886	AAQ03385	Aax03385 DNA seque
7	1413	98.7	886	AAI12393	Aai12393 Chinese c
8	1413	98.7	1227	AAQ06343	Aaq06343 Encodes T

9	1413	98.7	1227	2	AAQ31826	Aaq31826 Recombina
10	1413	98.7	1228	2	AAQ26499	Aaq26499 Trichosan
11	1411	98.5	870	6	ABA95171	Aba95171 Native tr
12	1397	97.6	885	2	AAQ37679	Aaq37679 Encode c
13	1326	92.6	1185	2	AAQ06351	Aaq06351 Alpha-Tri
14	1326	92.6	1185	2	AAQ06351	Aaq06351 Alpha-Tri
15	1214	84.8	800	2	AAQ06346	Aaq06346 Ribosome
16	1214	84.8	800	2	AAQ06346	Aaq06346 Ribosome
17	1211	84.6	751	2	AAQ26500	Aaq26500 Synthetic
18	1164.5	81.3	1094	2	AAQ75305	Aaq75305 Trichosan
19	1074	75.0	1499	4	AAT16208	Aat16208 Bryodin 1
20	1064	74.3	758	4	AAC60429	Aac60429 BDI-628.5
21	802.5	56.0	1026	2	AAC60425	Aac60425 DNA encod
22	798	55.7	1077	2	AAQ32687	Aaq32687 Luffin-f
23	798	55.7	1077	2	AAQ06348	Aaq06348 Alpha-Tri
24	798	55.7	1179	2	AAQ26502	Aaq26502 Ribosome
25	798	55.7	1180	2	AAQ06349	Aaq06349 Alpha-Tri
26	789.5	55.1	1169	2	AAQ26503	Aaq26503 Ribosome
27	789.5	55.1	1169	2	AAQ06350	Aaq06350 Ribosome
28	775	54.1	914	2	AAQ26504	Aaq26504 Alpha-Tri
29	766.5	53.5	1012	2	AAQ32688	Aaq32688 Ribosome
30	728	50.8	913	2	AAQ66093	Aaq66093 Luffin-g
31	728	50.8	913	2	AAQ42307	Aaq42307 Momordin
32	549	38.3	686	2	AAQ76184	Aaq76184 Type I ri
33	546	38.1	686	2	AAQ26501	Aaq26501 Ribosome
34	537.5	37.5	962	2	AAQ06347	Aaq06347 Trichosan
35	427	29.8	1855	2	AAQ89874	Aaq89874 Bryodin-2
36	427	29.8	1879	2	AAQ04222	Aaq04222 PAP-242 i
37	426	29.7	1923	1	AAN91039	Aan91039 Plasmid p
38	424.5	29.6	7306	7	ABZ77370	Abz77370 DNA seque
39	424.5	29.6	8155	7	ABZ77369	Abz77369 Nucleotid
40	423	29.5	1695	1	AAI64137	Aai64137 Preproric
41	423	29.5	1698	4	AAI64137	Aai64137 Castor be
42	423	29.5	1805	4	AAI64145	Aai64145 Modified
43	423	29.5	1855	2	AAT97913	Aat97913 Plasmid p
44	423	29.5	1855	2	AAT97919	Aat97919 Plasmid p
45	423	29.5	1855	2	AAT97910	Aat97910 Plasmid p

ALIGNMENTS

RESULT 1  
AAH41473  
ID AAH41473 standard; DNA; 870 BP.  
AC AAH41473;  
XX  
DT 23-AUG-2001 (first entry)  
XX  
DE Trichosanthes kirilowii trichosanthin (TCS) nucleotide sequence.  
XX  
KW Trichosanthes kirilowii; trichosanthin; TCS; mutagenesis; mutation; MTCS;  
KW mutant of trichosanthin; Mongolian snake-gourd; bioactivity; selectivity;  
KW cancer; virus; HIV; metaphase induced labour; ds.  
XX  
OS Trichosanthes kirilowii.  
XX  
PH Key Location/Qualifiers  
FT CDS 1..870  
FT /\*tag= a  
FT /product= "trichosanthin (TCS)"  
FT /transl\_except= (pos:688..690,aa:Gly)  
XX  
PN CN1283630-A.  
XX  
PD 14-FEB-2001.  
XX  
PF 02-AUG-2000; 2000CN-00119553.  
XX  
PR 02-AUG-2000; 2000CN-00119553.  
XX  
PA (SHAN-) SHANGHAI INST CYTOBIOLOGY CHINESE ACAD.  
XX

PI	Ke Y, Nie H;
XX	WPI; 2001-291745/31.
DR	P-PSDB; AAB99329.
XX	Trichosanthin mutant and its preparing process.
PT	Disclosure; Page 2 (disclosure); 15pp; Chinese.
XX	The present invention describes a trichosanthin mutant which is prepared through the mutational deformation of the trichosanthin gene and using an expression system. The trichosanthin gene is isolated from Trichosanthes kirilowii (Mongolian snake-gourd). The trichosanthin mutant has several advantages including high bioactivity, high selectivity to target, and strong kill action to cancer cells, virus and HIV. It can also be used for metaphase induced labour. The present sequence represents the nucleotide sequence of wild type trichosanthin which is given in the exemplification of the present invention
CC	Sequence 870 BP; 248 A; 193 C; 176 G; 253 T; 0 U; 0 Other;
SQ	
Alignment Scores:	
Pred. No.:	8.53e-142 Length: 870
Score:	1416.00 Matches: 286
Percent Similarity:	99.31% Conservative: 1
Best Local Similarity:	98.96% Mismatches: 2
Query Match:	98.88% Indels: 0
DB:	4 Gaps: 0
US-09-905-247-1 (1-289) x AAH41473 (1-870)	
QY	1 MetIleArgPheLeuValLeuSerLeuLeuLeuThrLeuPheLeuThrProAla 20
DB	1 ATGATCAGATTCTTAGTCCTCTTTTGGTCAATCTCACCTCTCTTAACAACTCGTGCT 60
QY	21 ValGluGlyAspValSerPheArgLeuSerGlyAlaThrSerSerSerTyrGlyValPhe 40
DB	61 GTGGAGGGCGATGTAGCTTCGGTTTATCAGGTGCACAACAGCAGTTCCTATGAGGTTTTTC 120
QY	41 IleSerAsnLeuArgLysAlaLeuProAsnGluArgLysLeuTyrAspIleProLeuLeu 60
DB	121 ATTTCAAATCTGAGAAAAGCTCTTCCAATAAAGAAGGAACCTGACCATATCCCTCTGTTA 180
QY	61 ArgSerSerLeuProGlySerGlnArgTyrAlaLeuLeuHisLeuThrAsnTyrAlaAsp 80
DB	181 CGTTCAGTCTTCCAGGTCTCAAGCTACGCAATGATCCATCTCACAAAATTACGCCGAT 240
QY	81 GluThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGly 100
DB	241 GAAACCATTTCACTGGCCATAGACGTACCAACGCTCTATATTATGGGATATCGCGTGGC 300
QY	101 AspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPheLys 120
DB	301 GATACATCCTATTTTCAACAGAGCTTCTCCAACAGAGCTGCAAAATATGATTTCAA 360
QY	121 AspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAla 140
DB	361 GACGCTATCGAAAAAGTTACGCTTCCATATTCTGGCAATTACGAAAGCTTCAAACTGCT 420
QY	141 AlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThr 160
DB	421 GCAGGCAAAATAAGGAAAAATATTCGGCTGGACTCCCTGCTTTGGACAGTGCATTAACC 480
QY	161 ThrLeuPheTyrTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSer 180
DB	481 ACTTTGTTTTACTACAACCCAAATTCGCTGGCTGGCACTTATGTTACTATTCATTCAGTCG 540
QY	181 ThrSerGluAlaAlaArgTyrLysPheIleGluGlnGlnIleGlyLysArgValAspLys 200
DB	541 ACCTCTGAGGCTGCGAGGTATAAATTTATTAGCAACAAATTTGGGAAGCGGTGTGACAAA 600
QY	201 ThrPheLeuProSerLeuAlaIleIleSerLeuGluAsnSerTrpSerAlaLeuSerLys 220





US-09-905-247-1 (1-289) x AAT61376 (1-886)

QY 1 MetIleArgPheLeuValLeuSerLeuLeuLeuThrLeuPheLeuThrThrProAla 20  
 Db 8 ATGATCAGATTCTTAGTCTCTCTTTGCTAATCTCACCCTCTCTCTTAAACACTCTGCT 67  
 QY 21 ValGluGlyAspValSerPheArgLeuSerGlyAlaThrSerSerSerGlyValPhe 40  
 Db 68 GTGAGGGCGATGTTAGCTTCGCTTTATCAGGTGCAACAGCAGTTCCTATGAGTTTC 127  
 QY 41 IleSerAsnLeuArgLysAlaLeuProAsnGluArgLysLeuTyrAspIleProLeuLeu 60  
 Db 128 ATTTCAAATCTGAGAAAGCTCTTCCAAATGAAGGAACTGTACGATATCCCTCTGTGA 187  
 QY 61 ArgSerSerLeuProGlySerGlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAsp 80  
 Db 188 CGTTCCTCTCTCCAGGTTCTCAACGCTACGCTATGATCCATCTCAAAATACGCCGAT 247  
 QY 81 GluThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGly 100  
 Db 248 GAAACCATTTTCAGTGCCATAGACGTAAACGCTCTATATATGCGATATCGCGTGC 307  
 QY 101 AspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaLysTyrValPheLys 120  
 Db 308 GATACATCTATTTTTCACGAGGCTTCTGCACAGAGCTGCAGAAATATGTATTCAAA 367  
 QY 121 AspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAla 140  
 Db 368 GAGCGTATGCGAAAGTTACGCTTCCATATCTGGCAATTACGAAGGCTTCAACTGCT 427  
 QY 141 AlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThr 160  
 Db 428 GCGGCAAAATAAGGAAATAATTCGCTTGACTCCAGCTTTGGACAGTCCATTACC 487  
 QY 161 ThrLeuPheTyrTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSer 180  
 Db 488 ACTTGTGTTTACTACAGCCCAATTCCTGCTGCGGCACCTTATGCTACTCAATTCAGTCG 547  
 QY 181 ThrSerGluAlaAlaArgTyrLysPheIleGluGlnGlnIleGlyLysArgValAspLys 200  
 Db 548 ACCTCTGAGGCTGCGAGGTATAAATTTATTGAGCAACAANTGGAGCGGCTTGACAAA 607  
 QY 201 ThrPheLeuProSerLeuAlaIleSerLeuGluAsnSerTyrSerAlaLeuSerLys 220  
 Db 608 ACCCTCTTACCAGTTAGCAATTAATAGTTTGGAAATAGTTGCTCTCTCTCCAAG 667  
 QY 221 GlnIleGlnIleAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuIleAsn 240  
 Db 668 CAAATTCAGATAGCGAGTACTTAATGAGCAGTTTGAACTCTCTGTTGTCTTATAAT 727  
 QY 241 AlaGlnAsnGlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnIle 260  
 Db 728 GCTCAAAACCAACGAGTCATGATAACCAATGTTGATGCTGAGTTGTAACTCCCAATC 787  
 QY 261 AlaLeuLeuLeuAsnArgAsnAsnMetAlaAlaMetAspAspValProMetThrGln 280  
 Db 788 CGGTGCTGCTGAACTGAAACCAATATGGACCAATGATGATGATGATGATGATGATGAT 847  
 QY 281 SerPheGlyCysGlySerTyrAlaLeu 289  
 Db 848 AGCTTTGGATGGAGTTATGCTATT 874

## RESULT 5

AAAX29143  
 ID AAAX29143 standard; DNA; 886 BP.

XX AC  
 AC AAAX29143;

DT 20-MAR-2003 (revised)

DT 04-JUN-1999 (first entry)

XX Chinese cucumber alpha-trichosanthin encoding DNA.  
 DE

XX KW Recombinant; plant virus; coat protein; systemic infection;  
 KW transcription; therapeutic; chinese cucumber; alpha-trichosanthin; ss.  
 XX OS Cucumis sp.  
 XX PN US5889190-A.  
 XX PD 30-MAR-1999.  
 XX PF 07-JUN-1995; 95US-00480432.  
 XX PR 26-FEB-1988; 88US-00160766.  
 PR 26-FEB-1988; 88US-00160771.  
 PR 15-JUL-1988; 88US-00219279.  
 PR 17-FEB-1989; 89US-00310881.  
 PR 05-MAY-1989; 89US-00347637.  
 PR 08-JUN-1989; 89US-00363138.  
 PR 22-OCT-1990; 90US-00600244.  
 PR 16-JAN-1991; 91US-00641617.  
 PR 26-JUL-1991; 91US-00737899.  
 PR 01-AUG-1991; 91US-00739143.  
 PR 31-JUL-1992; 92US-00923692.  
 PR 19-JAN-1994; 94US-00184237.  
 XX (BIOS-) BIOSOURCE TECHNOLOGIES INC.

XX Garger SJ, Turpen AM, Grill LK, Grantham GL, Donson J, Dawson WO;  
 PI Turpen TH;

XX WPI; 1999-243290/20.  
 DR P-PSDB; AAY01374.

XX Recombinant plant viral nucleic acid derived from a plus sense, single  
 stranded RNA plant virus - useful for the transcription of products in a  
 host.  
 XX Example 4; Col 43-46; 46pp; English.

XX The invention relates to a recombinant plant viral nucleic acid derived  
 from a plus sense, single stranded RNA plant virus. The recombinant plant  
 viral nucleic acid comprises: (a) a first plant viral subgenomic promoter  
 that is native to the plus sense, single stranded RNA plant virus and  
 operably joined to a first nucleic acid expression sequence; and (b) a  
 second plant viral subgenomic promoter that is non-native and is operably  
 joined to a second nucleic acid expression sequence; where, (i) (a) and  
 (b) are incapable of recombination with one another; (ii) either the  
 first or the second nucleic acid expression sequence is a plant viral  
 coat protein coding sequence. The recombinant plant viral nucleic acid  
 allows the transcription of products in a host, such as therapeutic and  
 other useful polypeptides or proteins e.g. enzymes, complex biomolecules  
 and ribozymes. It also gives the option of applying the coding sequence  
 to the desired organism, tissue, organ or cell, is stable for the foreign  
 coding sequences and is capable of systemic infection in the plant host.  
 CC The transformation and regeneration of target organisms become  
 CC unnecessary. (Updated on 20-MAR-2003 to correct PR field.)

XX Sequence 886 BP; 250 A; 199 C; 181 G; 256 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 1-83e-141 Length: 886  
 Score: 1413.00 Matches: 285  
 Percent Similarity: 99.31% Conservative: 2  
 Best Local Similarity: 98.62% Mismatches: 2  
 Query Match: 98.67% Indels: 0  
 DB: 2 Gaps: 0

US-09-905-247-1 (1-289) x AAAX29143 (1-886)

QY 1 MetIleArgPheLeuValLeuSerLeuLeuLeuThrLeuPheLeuThrThrProAla 20  
 Db 8 ATGATCAGATTCTTAGTCTCTCTTTGCTAATCTCACCCTCTCTCTTAAACACTCTGCT 67

QY 21 ValGluGlyAspValSerPheArgLeuSerGlyAlaThrSerSerSerTyrGlyValPhe 40  
 Db 68 GTGGAGGGCGATGTTAGCTTCGTTTATCAGGTGCAACAGCAGTTCCTATGGAGTTTC 127  
 QY 41 IleSerAsnLeuArgLysAlaLeuProAsnGluArgLysLeuTyrAspLeuLeu 60  
 Db 128 ATTTCAAATCTGAGAAAGCTTCCAAATGAAAGAAAGCTGACGATATCCCTCTGTGA 187  
 QY 61 ArgSerSerLeuProGlySerGlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAsp 80  
 Db 188 CGTTCCTCTTCAGGTTCTCAACGCTACGCTATGATCATCTCACAAATTACGCGCAT 247  
 QY 81 GluThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGly 100  
 Db 248 GAAACCAATTCAGTGGCCATAGACGTAAACGAGCTCTATATTATGGATATCGCGCTGGC 307  
 QY 101 AspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPheLys 120  
 Db 308 GATACATCCTATTTTTTCAAGGAGCTTCTGCAACAGAGCTGCAAAATATGATTCAAA 367  
 QY 121 AspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAla 140  
 Db 368 GACGCTATGCGAAGTTACGCTTCCATATTCTGCAATACGAAAGCTTCAAACTGCT 427  
 QY 141 AlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThr 160  
 Db 428 GCGGCAAAATAAGGAAATATTCCGCTTGGACTCCAGCTTTGGACAGTGCCATTACC 487  
 QY 161 ThrLeuPheTyrTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSer 180  
 Db 488 ACTTGTGTTTACTACAGCCCAATTCCTGCTGCTGGCAGCTATGTTACTCATTCAGTCG 547  
 QY 181 ThrSerGluAlaAlaArgTyrLysPheIleGluGlnGlnIleGlyLysArgValAspLys 200  
 Db 548 ACGTCTGAGGCTGCGAGGTATAAATTTATTGAGCAACAAATTTGGGAAGCGGTTGACAAA 607  
 QY 201 ThrPheLeuProSerLeuAlaIleIleSerLeuGluAsnSerTrpSerAlaLeuSerLys 220  
 Db 608 ACCTCTTACCAAGTTTACGATTAATTAAGTTTGGAAATAGTTGCTGCTCTCTCCAG 667  
 QY 221 GlnIleGlnIleAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuIleAsn 240  
 Db 668 CAATTCAGATAGCGAGTACTAATATGACAGATTTGAACCTCTGCTGCTTATAAT 727  
 QY 241 AlaGlnAsnGlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnIle 260  
 Db 728 GCTCAAAACCAACGAGTCATGATACCAATGTTGATGCTGGAGTTGTAACTCCCAATC 787  
 QY 261 AlaLeuLeuLeuAsnArgAsnMetAlaAlaMetAspAspValProMetThrGln 280  
 Db 788 GCGTTGCTGCTGAATCGAAACAAATATGGCAGCCATGGATGACGATGTTCTATGACACAG 847  
 QY 281 SerPheGlyCysGlySerTyrAlaLeu 289  
 Db 848 ACCTTTGGATGCGAAAGTTATGCTATT 874

RESULT 6  
 AAX03385  
 ID AAX03385 standard; DNA; 886 BP.  
 XX  
 AC AAX03385;  
 XX  
 DT 27-AUG-2003 (revised)  
 DT 01-APR-1999 (first entry)  
 XX  
 DE DNA sequence encoding chinese cucumber alpha-trichosanthin.  
 XX  
 KW Chinese cucumber; alpha-trichosanthin; plant virus;  
 KW RNA plant virus promoter; systemic infection; foreign gene expression;  
 KW AIDS therapeutic drug; ss.  
 XX  
 OS Momordica cochinchinensis.  
 XX

Location/Qualifiers  
 8..877  
 /\*tag= a  
 /product= "alpha-trichosanthin"  
 US5866785-A.  
 02-FEB-1999.  
 07-JUN-1995.  
 26-FEB-1988;  
 88US-00160766.  
 88US-00160771.  
 88US-00219279.  
 89US-00310881.  
 89US-00347637.  
 89US-00363138.  
 90US-00600244.  
 91US-00641617.  
 91US-00737899.  
 91US-00739143.  
 92US-00923692.  
 94US-00184237.  
 (BIOS-) BIOSOURCE TECHNOLOGIES INC.  
 Garger SJ, Turpen AM, Grill LK, Grantham GL, Donson J, Dawson WO;  
 Turpen TH;  
 WPI; 1999-142035/12.  
 P-PSDB; AAW84292.  
 Recombinant plant viral vector - that is capable of systemic infection in  
 host plant and stable production of heterologous DNA useful for producing  
 therapeutic proteins for treating e.g. AIDS.  
 Example 4; Col 43-46; 45pp; English.  
 The present sequence encodes chinese cucumber alpha-trichosanthin. The  
 nucleic acid sequence can be expressed in the plant viral constructs of  
 the invention. The specification describes a recombinant plant viral  
 nucleic acid derived from a positive (+)-sense RNA plant virus comprising  
 a native (+)-sense RNA plant virus promoter that is linked to an  
 expression sequence and a heterologous (+)-sense RNA plant virus promoter  
 that is linked to an expression sequence. The promoters are incapable of  
 recombination with each other, and one of the expression sequences  
 encodes a plant viral coat protein while the other is optionally a  
 heterologous coding sequence. The plant viral nucleic acid is capable of  
 systemic infection in a host plant. The viral construct is useful for the  
 introduction and expression of non-viral foreign genes in plants and the  
 production of e.g. potential AIDS therapeutic drugs. (Updated on 27-AUG-  
 2003 to correct OS field.)  
 Sequence 886 BP; 250 A; 199 C; 181 G; 256 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 1.83e-141 Length: 886  
 Score: 1413.00 Matches: 285  
 Percent Similarity: 99.31% Conservative: 2  
 Best Local Similarity: 98.62% Mismatches: 2  
 Query Match: 98.67% Indels: 0  
 DB: 2 Gaps: 0  
 US-09-905-247-1 (1-289) x AAX03385 (1-886)  
 QY 1 MetIleArgPheLeuValLeuSerLeuLeuIleLeuThrLeuPheLeuThrProAla 20  
 Db 8 ATGATCAGATTCCTAGTCCCTCTCTTGGCTAAATCTCACCCCTCTTCTACAACTCTGCT 67  
 QY 21 ValGluGlyAspValSerPheArgLeuSerGlyAlaThrSerSerSerTyrGlyValPhe 40  
 Db 68 GTGGAGGGCGATGTTAGCTTCGTTTATCAGGTGCAACAGCAGTTCCTATGGAGTTTC 127



QY 41 IleserAsnLeuArgLysAlaLeuProAsnGluArgLysLeuTyrAspIleProLeuLeu 60  
 Db 128 ATTTCAAACTGAGAAAGCTCTTCCAAATGAAAGGAAACTGTACGATATCCCTCTGTTA 187  
 QY 61 ArgSerSerLeuProGlySerGlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAsp 80  
 Db 188 CGTTCCTCTCTCCAGGTCTCAACGCTGCGCATTCATCCAAATACGCGCAT 247  
 QY 81 GluThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGly 100  
 Db 248 GAAACCATTTTCAGTGCCATAGACGTAAAGCAAGCTCTATATATGGGATATCCGCTGGC 307  
 QY 101 AspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaLysTyrValPheLys 120  
 Db 308 GATACATCTCTATTTTTCACGAGGCTTCGCAACAGAGCTGCAAAATATGATTCAA 367  
 QY 121 AspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAla 140  
 Db 368 GACGCTATCGAAAGTTACGCTTCATATTCGGCAATACGAAAGGCTTCAAACCTGCT 427  
 QY 141 AlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThr 160  
 Db 428 GCGGGCAAAATAGGGAATAATTCGCTTGGACTCCCAAGCTTGGACAGTGCATATACC 487  
 QY 161 ThrLeuPheTyrTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSer 180  
 Db 488 ACTTTGTTTACTACAGCCCAATTCGTGCTGGCTGCGCACTTATGCTACTCATTCAGTCG 547  
 QY 181 ThrSerGluAlaAlaArgTyrLysPheIleGluGlnGlnIleGlyLysArgValAspLys 200  
 Db 548 AGCTCTGAGGCTGCGAGGTATTAATTTATTTAGCAACAATTTGGGAAGCGCTTGACAAA 607  
 QY 201 ThrPheLeuProSerLeuAlaIleIleSerLeuGluAsnSerTyrSerAlaLeuSerLys 220  
 Db 608 ACCTTCTTACCAAGTTAGCAATTAATAGTTTGGAAATAGTTGCTGCTCTCTCCAG 667  
 QY 221 GlnIleGlnIleAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuIleAsn 240  
 Db 668 CAATTCAGATAGCGAGTACTAATTAATGACAGATTTGAAACTCTCTGTTGCTTATAAT 727  
 QY 241 AlaGlnAsnGlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnIle 260  
 Db 728 GCTCAAAACCAACAGTCATGATACCAATGTTGATGCTGGAGTTGTAACTCTCCAAATC 787  
 QY 261 AlaLeuLeuLeuAsnArgAsnAsnMetAlaAlaMetAspAspValProMetThrGln 280  
 Db 788 GCGTTGCTGCTGAATCGAAACAATATGCGCCATGCGATGATGCTTCTTATGACACAG 847  
 QY 281 SerPheGlyCysGlySerTyrAlaLeu 289  
 Db 848 AGCTTGGATGGGAAGTTATGCTATT 874

## RESULT 7

AAAL2393

ID AAA12393 standard; DNA; 886 BP.

XX AC AAA12393;

XX AC AAA12393;

DT 06-AUG-2003 (revised)

DT 24-AUG-2000 (first entry)

XX XX

XX XX

XX Chinese cucumber alpha-trichosanthin DNA.

XX Animal RNA virus; viral coat protein; plant; male sterility; interleukin;

XX EPO; erythropoietin; CSF; colony stimulating factor; Factor VIII; hGH;

XX human growth hormone; melanin; insulin; vaccine;

XX stereo specific catalysis; alpha-trichosanthin; ss.

XX Trichosanthin kirilowii.

OS XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

/product= "Alpha-trichosanthin"

US6054566-A.

25-APR-2000.

07-JUN-1995;

95US-00484341.

26-FEB-1988;

88US-00160766.

26-FEB-1988;

88US-00160771.

15-JUL-1988;

89US-00219279.

17-FEB-1989;

89US-00310881.

05-MAY-1989;

89US-00347637.

08-JUN-1989;

89US-00363138.

22-OCT-1990;

90US-00600244.

16-JAN-1991;

91US-00641617.

26-JUL-1991;

91US-00737899.

01-AUG-1991;

91US-00739143.

31-JUL-1992;

92US-00923692.

(BIOS-) BIOSOURCE TECHNOLOGIES INC.

Garger SJ, Grill LK, Turpen TH, Grantham GL, Dawson WO,

Turpen AM, Donson J;

WPI: 2000-338510/29.

P-PSDB; AAY87791.

Recombinant non-retroviral nucleic acid for producing proteins such as

interleukins, melanin and vaccines, comprises subgenomic promoters linked

to sequences coding for viral coat protein and heterologous proteins.

Example 4; Col 57-60; 51pp; English.

This invention describes a novel recombinant viral nucleic acid (I) from

a non-retroviral (+) sense, single stranded animal RNA virus comprising a

nucleic acid sequence coding for a viral coat protein regulated by a

native subgenomic promoter and other two heterologous nucleic acid

sequences regulated by two other subgenomic promoters. (I) is useful for

expressing foreign genes e.g. genes inducing male sterility in plants.

(I) is also useful for producing proteins such as interleukins, EPO

(erythropoietin), CSF (colony stimulating factor), Factor VIII, hGH

(human growth hormone), melanin, insulin, vaccines etc., and enzymes that

are useful for stereo specific catalysis of organic compounds. (I) is

stable and transcribed systemically. The dual subgenomic promoter system

reduces the frequency of recombination thus reducing regeneration of the

wild type virus. This sequence encodes a chinese cucumber alpha-

trichosanthin protein which is described in the method of the invention.

(Updated on 06-AUG-2003 to correct OS field.)

Sequence 886 BP; 250 A; 199 C; 181 G; 256 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.83e-141 Length: 886

Score: 1413.00 Matches: 285

Percent Similarity: 99.31% Conservative: 2

Best Local Similarity: 98.62% Mismatches: 2

Query Match: 98.67% Indels: 0

DB: 3 Gaps: 0

US-09-905-247-1 (1-289) x AAA12393 (1-886)

QY 1 MetIleArgPheLeuValLeuSerLeuLeuIleLeuThrLeuPheLeuThrProAla 20

Db 8 ATGATCAGATTCCTAGTCTCTCTTGTCTAATCTCACCTCTCTCTAACTCTCTGCT 67

QY 21 ValGluGlyAspValSerPheArgLeuSerGlyAlaThrSerSerSerTyrGlyValPhe 40

Db 68 GTGGAGGGCGATGTTAGCTTCGCTTTATCAGGTGCAACAGCAGTTCCTATGAGTTTC 127

QY 41 IleSerAsnLeuArgLysAlaLeuProAsnGluArgLysLeuTyrAspIleProLeuLeu 60

Db 128 ATTTCAAACTGAGAAAGCTCTTCCAAATGAAAGGAAACTGTACGATATCCCTCTGTTA 187

QY	61	ArgSerSerLeuProGlySerGlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAsp	80
Db	188	CGTTCCTCTCTCTCCAGGTTCCTCAACGCTAGCATTGATCCATCTCAACAATTACGCCGAT	247
QY	81	GluThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGly	100
Db	248	GAACACATTTTCAGTGCCATAGACGTAAACAAAGCTCTATATTATGGGATATCGCGTGGC	307
QY	101	AspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPheLys	120
Db	308	GATACATCTCTATTTTTCACAGAGGCTTCGCAACAGAGCTGCAAAATATGATTCAAA	367
QY	121	AspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAla	140
Db	368	GAGCTATGCAAAAGTTACGCTTCCTATATCTGGCAATTACGAAAGCTTCAAACTGCT	427
QY	141	AlaGlyLysIleArgGluAunIleProLeuGlyLeuProAlaIleuAspSerAlaIleThr	160
Db	428	GCGGGCAAAATAAGGGAAAAATATTCGCTTGGACTCCGAGCTTTGGAGACGTGCCATTACC	487
QY	161	ThrLeuPheTyrTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSer	180
Db	488	ACTTTGTTTCTACTACACGCCAATTCGTGCGTCCGACATTATGGTACTCATTCAGTCG	547
QY	181	ThrSerGluAlaAlaArgTyrLysPheIleGluGlnIleGlyLysArgValAspLys	200
Db	548	ACGTCCTGAGGCTCGAGGTAATAATTTATTGAGCAACAAATTTGGAAAGCGGTTGCACAA	607
QY	201	ThrPheLeuProSerLeuAlaIleIleSerLeuGluAsnSerTrpSerAlaLeuSerLys	220
Db	608	ACCTTCCTACCAAGTTTAGCAATATATAAGTTTGGAAATATGTTGGTCTGCTCTCTCCAG	667
QY	221	GlnIleGlnIleAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuIleAsn	240
Db	668	CAAAATTCAGATAGCGAGTACTAATAATGACAGATTTGAAACTCCTCTGTTGTGCTTATAAAT	727
QY	241	AlaGlnAsnGlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnIle	260
Db	728	GCTCAAAACCAACGAGTCATGATAACCAATGTTGATGCTGGAGTTGTAACTCCCAACATC	787
QY	261	AlaLeuLeuLeuAsnArgAsnAsnMetAlaAlaMetAspAspValProMetThrGln	280
Db	788	CGCTTGCTGCTGAATCGAAACAATATGGCAGCATGGATGACGATGTTCCCTATGACACAG	847
QY	281	SerPheGlyCysGlySerTyrAlaLeu	289
Db	848	TCCTTTTGGTGTGGAGTTATGCTATT	874

## RESULT 8

res001	
AAQ06343	
ID	AAQ06343 standard; DNA; 1227 BP.
XX	
AC	AAQ06343;
XX	
DT	06-FEB-1991 (first entry)
XX	
DE	Encodes Trichosanthin from Trichosanthus kirilowii.
XX	
KW	trichosanthin (TCS); Human Immunodeficiency Virus; inhibition; ss.
XX	
OS	Trichosanthus kirilowii.
XX	
FH	Key Location/Qualifiers
CDS	340..1209
FT	/*tag= a
FT	409..1209
FT	/*tag= b
FT	/product= "TCS"
XX	
PN	WO9012097-A.
XX	
PD	18-OCT-1990.

[illegible]





ID ABA95171 standard; cDNA; 870 BP.  
 XX AC ABA95171;  
 XX DT 20-MAY-2002 (first entry)  
 XX DE Native trichosanthin (TCS) encoding cDNA.  
 XX KW Trichosanthin; TCS; mutant; MTCS; therapeutic; cytostatic; antitumour;  
 XX KW anti-human immunodeficiency virus; virucide; immunostimulant; gene;  
 XX KW ectopic pregnancy; ss.  
 XX OS Trichosanthes kirilowii.  
 XX FH Key Location/Qualifiers  
 XX FT CDS 1..870  
 XX FT /tag= a  
 XX FT /transl\_except= (pos: 16..18, aa: Val)  
 XX FT /transl\_except= (pos: 523..525, aa: Met)  
 XX FT /transl\_except= (pos: 865..867, aa: Leu)  
 XX FT /product= "TCS"  
 XX FT sig\_peptide 1..69  
 XX FT mat\_peptide 70..867  
 XX FT /tag= b  
 XX FT /tag= c  
 XX PN WO200212537-A2.  
 XX PD 14-FEB-2002.  
 XX PF 18-JUL-2001; 2001WO-CN001178.  
 XX PR 02-AUG-2000; 2000CN-00119553.  
 XX PR 18-JAN-2001; 2001CN-00103102.  
 XX PA (BEIJ-) BEIJING STM BIOTECH LTD.  
 XX PA (KEY/) KE Y.  
 XX PA (NIEH/) NIE H.  
 XX PI Ke Y, Nie H;  
 XX DR WPI; 2002-227165/28.  
 XX DR P-PSDB; ABB07660.  
 XX PT Mutant trichosanthin protein of low antigenicity useful for treating  
 XX PT tumor e.g. leukemia, comprises a sequence of native trichosanthin with  
 XX PT modification of at least one amino acid residue in three specific  
 XX PT regions.  
 XX PS Example 1; Fig 1; 42pp; English.  
 XX CC The invention relates to a mutant trichosanthin (MTCS) protein of low  
 XX CC antigenicity comprising a sequence of native TCS with the modification of  
 XX CC at least one amino acid residue in three regions which is 174-180, 203-  
 XX CC 227 and 230-244 and substantially retaining the biological activities of  
 XX CC the native TCS. The MTCS protein, its fragment or derivative is useful as  
 XX CC an therapeutic agent; or for preparing a medicament for treating viral  
 XX CC disease e.g. acquired immunodeficiency syndrome (AIDS), tumour e.g.  
 XX CC leukemia, for inducing abortion and/or treating ectopic pregnancy. The  
 XX CC present sequence represents the native TCS encoding cDNA  
 XX SQ Sequence 870 BP; 249 A; 192 C; 176 G; 253 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 2.92e-141 Length: 870  
 Score: 1411.00 Matches: 285  
 Percent Similarity: 99.31% Conservative: 2  
 Best Local Similarity: 98.62% Mismatches: 2  
 Query Match: 98.53% Indels: 0  
 DB: 6 Gaps: 0  
 US-09-905-247-1 (1-289) x ABA95171 (1-870)

US-09-905-247-1 (1-289) x ABA95171 (1-870)

Recombinant products; commercial production; fermentation; biosynthesis;



PI	Platek M, Chow T, Fry K;
XX	WPI; 1990-334847/44.
DR	P-PSDB; AAR07523.
XX	
PT	Recombinant tri:chosanthin protein - with selective inhibitory effect on
PT	viral expression in HIV infected T-cells or monocyte-macrophase.
XX	
PS	Example; Fig 20; 102pp; English.
XX	
CC	pQ30E was used as a probe to identify clones containing sequences coding
CC	for TCS in a T.kirilowii genomic library. Five clones were eventually
CC	isolated and sequenced, including clone 12. They were found to have
CC	homology to the alpha-TCS coding sequence. See also AAQ06343-Q06350
XX	
SQ	Sequence 1185 BP; 372 A; 249 C; 221 G; 343 T; 0 U; 0 Other;

  

Alignment Scores:			
Pred. No.:	5.5e-132	Length:	1185
Score:	1326.00	Matches:	267
Percent Similarity:	95.16%	Conservative:	8
Best Local Similarity:	92.39%	Mismatches:	14
Query Match:	92.60%	Indels:	0
DB:	2	Gaps:	0

  

US-09-905-247-1 (1-289) x AAQ06351 (1-1185)

QY	1	MetlleArgPheLeuValLeuSerLeuLeuLeuThrLeuPheLeuThrProAla	20
Db	269	ATGATCAGATTCTTAGTCTTCCTTTGTCTAATCTCACCTCTTCCTAACAGCTCGTGCT	328
QY	21	ValGlUGlyAspValSerPheArgLeuSerGlyAlaThrSerSerSerTyrglyValPhe	40
Db	329	GTCCAGGGCGATGTTAGCTTCGGTTTATCAGGTGCACCAAGCAGTTCCCTATGGAGTTTC	388
QY	41	IleSerAsnLeuAcgLyseAlaLeuProAsnGluArgLysLeuTyrrAspIleProLeuLeu	60
Db	389	ATTTCAAATATGAGAAAGACTCTTCATATGAAGAGAGACTATACCAGATATCTCTGTTA	448
QY	61	ArgSerSerLeuProGlySerGlnArgTyraLaLeuLeuIleHisLeuThrAsnTyraAsp	80
Db	449	CGTTCACCTCTTCAGGTCTCAAGCTTCAGCTACGATGATCATCTCAAAATATACGCCGAT	508
QY	81	GluThrIleSerValAlalaileAspValThrAsnValTyriMetGlyTyraArgAlaGly	100
Db	509	GAACACCATTTGAGTGGCCATAGACGTAAACGACGCTATGTTATGGATATCGCGTGTG	568
QY	101	AspThrSerTyrrPhePheAsnGlyAlaSerAlaThrGluAlaAlaLysTyrrValPheLys	120
Db	569	GATACATCTCTATTTTTTCAACGAGGCTTCGCACTGAAAGCTGCATAATATGTTATTTCAA	628
QY	121	AspAlaMetArgLysValThrLeuProTyrrSerGlyAsnTyrrGluArgGluGlnThrAla	140
Db	629	GACGCTCAGCGAAAAGTTACGCTCCATATTCGCAATATGAAAGGCTTCAAAATGCA	688
QY	141	AlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThr	160
Db	689	GCAGGCAAAATAGGGAAAAATATTCGGCTTGAGCTACCTGCTTTGCAGAGTCCCATACC	748
QY	161	ThrLeuPheTyrrTyrrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeulleGlnSer	180
Db	749	AAYTTGTTTACTACAACGCCAAATTCCTCGCTCGGCACCTATGTTGTTACTATTTCAGTCG	808
QY	181	ThrSerGluAlaAlaArgTyrrLysPheIleGluGlnIleGlyLysArgValAspLys	200
Db	809	ATGTCCTGAGGCTGCGAGGTATAAATTTATTCAGCANACAAATTTGGAGCGGTGTGACAAA	868
QY	201	ThrPheLeuProSerLeuAlaIleIleSerLeuGluAsnSerTrpSerAlaLeuSerLys	220
Db	869	ACCTTCATCCCAAGTCTTGCAATATAAGTTTGGAAAATAGTTGCTGCTCTCTCAAG	928
QY	221	GlnIleGlnIleAlaSerThrAsnAsnClyGlnPheGluSerProValValLeuIleAsn	240

RESULT 15  
AAQ06346  
ID AAQ0  
XX  
AC AAQ0  
XX  
DT 06-F  
XX



```
Db 341 CGTAAAGTGACCCCTGCCATACAGCGGCAACTATGAACGTCGTGACAGCGGCGGTAAG 400
QY 144 IleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThrThrLeuPhe 163
Db 401 ATCCGTGAGAACATTCCGCTTGCTTACAGCTCTAGACTCCCGATCAGACTCTCTTC 460
QY 164 TyrTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSerThrSerGlu 183
Db 461 TATTACAACGCCAAATCCGCGCAAGCGGCTCATGGTTTGGATTCAAAGTACTAGCGAA 520
QY 184 AlaAlaArgTyrIlyspheIleGluGlnGlnIleGlyIysArgValAspLysThrPheLeu 203
Db 521 GCTGCACGCTACAAATTCTCGAACAGCAAAATGGCAACGCGTGGACAAACCTTCTG 580
QY 204 ProSerLeuAlaIleIleSerLeuGluAsnSerTrpSerAlaLeuSerLysGlnIleGln 223
Db 581 CCTTCGCTGCCATTATCTCTCTTGAGAATTCGTGGTCTGCCCTGAGCAAGCAGATCCAG 640
QY 224 IleAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuIleAsnAlaGlnAsn 243
Db 641 ATCGCGTCGACCACCAATGTCAGTTCGAAAGCCCGTCGTTTAAATTAATGCGCAGAAT 700
QY 244 GlnArgValThrIleThrAsnValIleAlaGlyValValThrSerAsnIleAlaLeuLeu 263
Db 701 CAACGTGTGACCATCAGCAACGTGGATGCGGAGTTGTGACGTCCAATATCGCCCTGCTT 760
QY 264 LeuAsnArgAsnAsnMetAla 270
Db 761 TTAACCGTAACAATATGSCA 781
```

Search completed: September 14, 2004, 20:19:35  
Job time : 393 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 14, 2004, 19:57:42 ; Search time 2476 Seconds  
(without alignments)  
3485.529 Million cell updates/sec

Title: US-09-905-247-1  
Perfect score: 1432  
Sequence: 1 MIRFLVLSLLILFLTPA.....AAMDDVPMTQSGCGSYAL 289

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool/p/US09905247/runat\_10092004\_103934\_4726/app\_query.fasta\_1.455  
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0 -L-LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HRAFSIZE=500 -MINLEN=2000000000  
-USER=US09905247@cgn\_1\_1 5180 -runat\_10092004\_103934\_4726 -NCPU=6 -ICPU=3  
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_man:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	230.5	16.1	665	14	CF227084	CF227084 IHSLEP22 S
2	222	15.5	661	14	CF227046	CF227046 IH202 sub
3	220	15.4	837	9	AW053634	AW053634 L30-1401T
4	220	15.4	1021	10	BE036639	BE036639 MP03B03 M
5	214	14.9	662	14	CF227047	CF227047 IH215 sub
6	198.5	13.9	658	13	BQ583480	BQ583480 E011979-0
7	197.5	13.8	601	14	CA198032	CA198032 SCEZAD107
8	193	13.5	1033	10	BE036541	BE036541 MPO1B07 M
9	181	12.6	420	14	CF227009	CF227009 IH149 sub
10	181	12.6	993	10	BE035039	BE035039 MM02A03 M
11	177.5	12.4	1038	10	BE033546	BE033546 MFO3A09 M
12	171	11.9	698	14	CA838926	CA838926 MCT021H09
13	169.5	11.8	667	14	CA838446	CA838446 MCT016F04
14	169.5	11.8	678	14	CA840373	CA840373 MCT036G03
15	169.5	11.8	811	10	BE034055	BE034055 MGO4C05 M
16	166.5	11.6	644	14	CA838401	CA838401 MCT016B04
17	163.5	11.4	611	13	BQ490182	BQ490182 34-E9435-
18	156.5	10.9	661	14	CA835532	CA835532 MCS038A11
19	156.5	10.9	689	14	CA833333	CA833333 MCS021A09
20	156.5	10.9	689	14	CA833383	CA833383 MCS021F04
21	152.5	10.6	649	14	CA838757	CA838757 MCT020B02
22	152.5	10.6	649	14	CA833930	CA833930 MCT026C11
23	152.5	10.6	671	14	CA839511	CA839511 MCT028D04
24	150.5	10.5	639	14	CA835855	CA835855 MCS004G12
25	143.5	10.0	405	14	T24255	T24255 CFS1345 lam
26	140.5	9.8	1094	11	AY105813	AY105813 Zea mays
27	139.5	9.7	816	10	BE035038	BE035038 MM02A01 M
28	132	9.2	639	10	BE037217	BE037217 MP18B02 M
29	126.5	8.8	677	14	CA174183	CA174183 SCJFST100
30	126	8.8	560	14	CF974192	CF974192 PSU meup2
31	126	8.8	607	14	CA222363	CA222363 SCSGFL403
32	125	8.7	680	14	CA198271	CA198271 SCCCF1300
33	124	8.7	419	14	CF226983	CF226983 IH114 sub
34	124	8.7	422	14	CF227044	CF227044 IH198 sub
35	123	8.6	620	13	CA080068	CA080068 SCSGAM109
36	121.5	8.5	672	9	AW053585	AW053585 L30-2143T
37	121.5	8.5	680	13	CA076205	CA076205 SMCAM110
38	121	8.4	765	14	CA195191	CA195191 SCEZSB109
39	120.5	8.4	730	13	CA088699	CA088699 SCRFBAM212
40	118.5	8.3	498	13	BQ588134	BQ588134 E012337-0
41	118.5	8.3	603	13	BQ590856	BQ590856 E012599-0
42	118.5	8.3	660	14	CK113535	CK113535 UR108TF08
43	118.5	8.3	752	13	CA122188	CA122188 SCFURL107
44	118	8.2	713	12	BG367523	BG367523 HVSME1001
45	118	8.2	817	12	BG369108	BG369108 HVSME1002

# ALIGNMENTS

RESULT 1  
LOCUS CF227084  
DEFINITION IHSLEP22 subtracted Iris tepal cDNA library Iris hollandica CDNA  
ACCESSION CF227084  
VERSION CF227084.1 GI:33438122  
KEYWORDS EST.  
SOURCE Iris hollandica  
ORGANISM Iris hollandica  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Iridaceae;  
Iris.

REFERENCE 1 (bases 1 to 665)  
 AUTHORS van Doorn,W.G., Balk,P.A., van Houwelingen,A.M., Hoerichs,F.A.,  
 Hall,R.D., Vorst,O., van der Schoot,C. and van Wordragen,M.F.  
 TITLE Gene expression during anthesis and senescence in Iris flowers  
 JOURNAL Unpublished (2003)  
 COMMENT Contact: Wouter G. van Doorn  
 Institute for Agrotechnological Research (ATO)  
 Wageningen University and Research Centre  
 P.O. Box 17, 6700 AA, Wageningen, The Netherlands  
 Tel: 0031 317 475105  
 Fax: 0031 317 475347  
 Email: wouter.vandoorn@wur.nl.

FEATURES  
 source  
 1..665  
 /organism="Iris hollandica"  
 /mol\_type="mRNA"  
 /cultivar="Blue Magic"  
 /db\_xref="taxon:35876"  
 /clone="IHSLP22"  
 /tissue\_type="sepal"  
 /clone\_lib="subtracted Iris tepal cDNA library"  
 /note="Vector: pGEMT-Easy; mRNA isolated from Iris sepals  
 was subtracted using pedicel material as the driver. Equal  
 amounts of mRNA from each developmental stage were pooled  
 for cDNA synthesis. The pGEMT-Easy Vector System  
 (Promega) was used to construct the cDNA library."

ORIGIN  
 Alignment Scores:  
 Pred. No.: 4,2e-18 Length: 665  
 Score: 230.50 Matches: 71  
 Percent Similarity: 51.98% Conservativeness: 47  
 Best Local Similarity: 31.28% Mismatches: 86  
 Query Match: 16.10% Indels: 24  
 DB: 14 Gaps: 7

US-09-905-247-1 (1-289) x CF227084 (1-665)  
 QY 36 SerTyrGlyValPheIleSerAsnLeuArgLysAlaLeuProAsnGluArgLysLeuTyr 55  
 Db 6 ACCATATAGTCTTCTCCTACAGACCTTACGACCGCTTAAGCAGTGGGAAGTAGCGTGCAT 65  
 QY 56 AsplleProLeuLeuArgSerSerLeuProGlySerGlnArgTyrAlaLeuIleHisLeu 75  
 Db 66 GATATTCCTTCCTTCCTCGCAATCCGGTTCTCAGCAGGATTTCTCTTGTGCGACTC 125  
 QY 76 ThrAsnTyrAlaAspGluThrIleSerValAlaIleAspValThrAsnValTyrIleMet 95  
 Db 126 TTGACTGGCCACACAGCGCATCAGCTAGTGTAAATCGCGTCAATGCTTATGTGGTG 185  
 QY 96 GlyTyrArgAlaGlyAspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAla 115  
 Db 186 GCTTATCAAGCTCAAAATCGTTTACTGTCTTCCGGC-----ACTCCCGCAAT 236  
 QY 116 LysTyrValPheLysAspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGlu 135  
 Db 237 CCTCAGTTTATGGAACAACCCACACCGCTTACT-----TTTACAGTACGTATGCT 290  
 QY 136 ArgLeuGlnThrAlaAlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeu 155  
 Db 291 GCATTCAAATGTTGGAAACCAAGTAGAGAAATATTGATTGGGGAATCAATCCACTT 350  
 QY 156 AspSerAlaIleThrThrLeuPheTyrTyrAsnAlaAsn-----Ser 169  
 Db 351 GCAACGGCAATCTCGACACTC-----CACAACTCGTCTCCGCTACGGTGGAAACATCC 404  
 QY 170 AlaAlaSerAlaLeuMetValLeuIleGlnSerThrSerGluAlaAlaArgTyrLysPhe 189  
 Db 405 GTGGCTGTTTCCCTTATCTTTTATTCATTCATTTGTTCTCCGAGCCGACGATTCAGAGCG 464  
 QY 190 IleGluGlnGlnIleGlyLysArg---ValAspLys-----Thr 201  
 Db 465 ATCAGGACACGAGTTTACAAACACATCATGACACCGAGTCAACCCCAATCCGCTACGACAA 524

QY 202 PheLeuProSerLeuAlaIleLeuSerLeuGluAsnSerTrpSerAlaLeuSerLysGln 221  
 Db 525 TTTCGCCCTAGTGTGGCATAATAGACCTACACTAGTCAAACCCCTCCACTGAA 583  
 QY 222 IleGlnIleAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuIleAsnAla 241  
 Db 584 GTTCAGCGAGCC-----GAAGAGGTAGTGTTCAGCCAGTAAATTCAGAGNTTCC 637  
 QY 242 GlnAsnGlnArgValThrIle 248  
 Db 638 GTCCACCAACACAGTCGGTATC 658

RESULT 2  
 CF227046 661 bp mRNA linear EST 04-AUG-2003  
 LOCUS IH202 subtracted Iris tepal cDNA library Iris hollandica cDNA clone  
 DEFINITION IH202, mRNA sequence.  
 ACCESSION CF227046 GI:33438082  
 VERSION CF227046  
 KEYWORDS EST.  
 SOURCE Iris hollandica  
 ORGANISM Iris hollandica  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Iridaceae;  
 Iris.  
 REFERENCE 1 (bases 1 to 661)  
 AUTHORS van Doorn,W.G., Balk,P.A., van Houwelingen,A.M., Hoerichs,F.A.,  
 Hall,R.D., Vorst,O., van der Schoot,C. and van Wordragen,M.F.  
 TITLE Gene expression during anthesis and senescence in Iris flowers  
 JOURNAL Unpublished (2003)  
 COMMENT Contact: Wouter G. van Doorn  
 Institute for Agrotechnological Research (ATO)  
 Wageningen University and Research Centre  
 P.O. Box 17, 6700 AA, Wageningen, The Netherlands  
 Tel: 0031 317 475105  
 Fax: 0031 317 475347  
 Email: wouter.vandoorn@wur.nl.

FEATURES  
 source  
 1..661  
 /organism="Iris hollandica"  
 /mol\_type="mRNA"  
 /cultivar="Blue Magic"  
 /db\_xref="taxon:35876"  
 /clone="IH202"  
 /tissue\_type="sepal"  
 /clone\_lib="subtracted Iris tepal cDNA library"  
 /note="Vector: pGEMT-Easy; mRNA isolated from Iris sepals  
 was subtracted using pedicel material as the driver. Equal  
 amounts of mRNA from each developmental stage were pooled  
 for cDNA synthesis. The pGEMT-Easy Vector System  
 (Promega) was used to construct the cDNA library."

ORIGIN  
 Alignment Scores:  
 Pred. No.: 4,93e-17 Length: 661  
 Score: 222.00 Matches: 71  
 Percent Similarity: 50.67% Conservativeness: 43  
 Best Local Similarity: 31.56% Mismatches: 88  
 Query Match: 15.50% Indels: 24  
 DB: 14 Gaps: 7

US-09-905-247-1 (1-289) x CF227046 (1-661)  
 QY 1 MetIleArgPheLeuValLeuSer-----LeuLeuIleLeuThrLeuPheLeuThr 17  
 Db 7 ATGAACAGCTGGTTGATCTTGGCAGTGACGTGATTTATGGGTACGATTGCTGGACCT 66  
 QY 18 ThrProAlaValGluGlyAspValSerPheArgLeuSerGlyAlaThrSerSerTyr 37  
 Db 67 GTAGTCTTCATTGAA---ACAGTGCAGTTCCTGTCCTGGGACCCAGGCAACCTAT 123  
 QY 38 GlyValPheIleSerAsnLeuArgLysAlaLeuProAsnGluArgLysLeuTyrAspIle 57

```

Db      124 AGTGGTTTCTTACAGACCTCAGAACGCTTAAGCAGTGGAGTAGCGTGATGATATT 183
QY      58 ProLeuLeuArgSerSerLeuProGlySerGlnArgTyrAlaLeuIleHisLeuThrAsn 77
Db      184 CTTTGTGTCGCGGCAATCCGGTCCAGAGGATTGCTCTTGTGCAACTCTTTGAT 243
QY      78 TyrAlaAspGluThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyr 97
Db      244 TGGGCAACACCGCATCCGCTGGTCTTAATCGCGTCAATGCTTATCTGTAGCTTAT 303
QY      98 ArgAlaGlyAspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaAlaLysTyr 117
Db      304 CAAGCTAAATTCGTATTACTTGTT-----TCCGACACTCCGCCCAATCCTCAG 354
QY      118 ValPheLysAspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeu 137
Db      355 CTTTATGGAAGCAACCCGACCGC-----CTTAGTTTACAGGAGCTACATGCTCTT 408
QY      138 GlnThrAlaAlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSer 157
Db      409 CAAATGTTGGCAAAACAAGTAGAAAATATTGATCTGGGAATCAATCCACTTCGAACG 468
QY      158 AlaIleThrThrLeuPheTyr-----TyrAsnAlaAsnSerAlaAlaSerAla 173
Db      469 GCATCAGCACACTCCA-CAACTGGGCTCCGCTACGCTGGAGACATCCGTGGCTCGTCC 527
QY      174 LeuMetValLeuIleGlnSerThrSerGluAlaAlaArgTyrLysPheIleGluGln 193
Db      528 CTTATCGTCTGATCCAACTGCTCCGAGCAGCAGCATTCAGAGCATCGAGCAACGA 587
QY      194 Iledglyys-----ArgValAspLysThrPheLeuPro 204
Db      588 GTTACGAGCAGCATATAGACGCGCTACCCCTATCCGTTACGAC---AGCTTTCGCTCT 644
QY      205 SerLeuAlaIleIle 209
Db      645 GGTGTTGGCAATAA 659

RESULT 3
AW053634
LOCUS
DEFINITION
L30-140173 Ice plant Lambda Uni-Zap XR expression library, 30 hours
NaCl treatment Mesembryanthemum crystallinum cDNA clone L30-1401
5', mRNA sequence.
AW053634
AW053634.1 GI:5916827
EST.
Mesembryanthemum crystallinum (common iceplant)
Mesembryanthemum crystallinum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Caryophyllales; Aizoaceae; Mesembryanthemum.
1 (bases 1 to 837)
Cushman, J.C.
An expressed sequence tag database for the common ice plant,
Mesembryanthemum crystallinum
Unpublished (1997)
Contact: Cushman JC
Department of Biochemistry
University of Nevada
MS200, Reno, NV 89557-0014, USA
Tel: 775-784-1918
Fax: 775-784-1650
Email: jcushman@unr.edu
Poly(A) tail, 10 nt: 838. .847
PCR Primers
FORWARD: T7
BACKWARD: T3
Insert Length: 847 Std Error: 5.00
Plate: L30-15 row: A column: 5
Seq primer: T3
High quality sequence stop: 450.

FEATURES
Source
Location/Qualifiers
1..837
/organism="Mesembryanthemum crystallinum"
/mol_type="mRNA"
/db_xref="taxon:3544"
/clone="L30-1401"
/tissue_type="Leaf, 30 h 0.4M NaCl"
/dev_stage="Six week old"
/clone_lib="Ice plant Lambda Uni-Zap XR expression
library, 30 hours NaCl treatment"
/note="Vector: Lambda Uni-Zap XR, Bluescript SK-; Site_1:
EcoRI; Site_2: XhoI"

ORIGIN
Alignment Scores:
Pred. No.: 1,3e-16 Length: 837
Score: 220.00 Matches: 69
Percent Similarity: 51.20% Conservative: 38
Best Local Similarity: 33.01% Mismatches: 73
Query Match: 15.36% Indels: 30
DB: 9 Gaps: 8

US-09-905-247-1 (1-289) x AW053634 (1-837)
QY 35 SerSerTyrGlyValPheIleSerAsnLeuArgLysAlaLeuProAsnGluArgLysLeu 54
Db 5 AACACATATTCACATTCATGACATCTCTCGAGCTCAACTCTCTGGCACCAACCGCATGT 64
QY 55 TyrAspIleProLeuLeuArgSerSerLeuProGlySerGlnArgTyrAlaLeuIleHis 74
Db 65 ---CAATCCCGGCGACAGCGTCAACGCCAGCATCTCCAAACGATTCTGCTCGTCGAC 121
QY 75 LeuThrAsnTyrAlaAspGluThrIleSerValAlaIleAspValThrAsnValTyrIle 94
Db 122 CTCAAAACAACCTCGCAAAAGACCATCACATCTGCAATCGACGTGACAGCGTCTATGTC 181
QY 95 MetGlyTyrArg-----AlaGlyAspThrSerTyrPhePheAsnGlyAlaSer 110
Db 182 GTGGCTTACCGCGACAAGCTTGGCGGAAAGACGCTGCCCACTTCTTAGCGATGCT--- 238
QY 111 AlaThrGluAlaAlaLysTyrValPheLysAspAlaMetArgLysValThrLeuProTyr 130
Db 239 CCCACGGTCTGCTAGGAACAACCTCTTCAAGGGCGGAGGTTCGG---AACATTGCGTTC 295
QY 131 SerGlyAsnTyrGluArgLeuGlnThrAlaAlaGlyLysIleArgGluAsnIleProLeu 150
Db 296 GGAGGGAGTTACATAGCCTTAGAGAGGCTGCAAGCAAGCGCGGAAATGCAATAGACTTA 355
QY 151 GlyLeuProAlaLeuAspSerAlaIleThrThrLeuPhe-----TyrTyrAsnAla 167
Db 356 GGGGTGAATTAACATAGAGTTTGGATCGAGTCGGTCTTTTGTAGAACCCGATTAATGGG 415
QY 168 AsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSerThrSerGluAlaAlaArgTyr 187
Db 416 CAGGTTGAGGCCAAATCTTGTGATGCTATACAGATGGTTCTTGAAGCAGACGCGTTT 475
QY 188 LysPheIleGluGln-----GlnIleGlyLysArgValAspLysThrPheLeuPro 204
Db 476 AAGTATATTGAAGTAAGGGACCCCAAGTGGG-----TTACATGGTGGTGGTCAACACC 529
QY 205 SerLeuAlaIleIleSerLeuGluAsnSerTrpSerAlaLeuSerLysGlnIleGlnIle 224
Db 530 GACCC- GAAGTCTGAGTTGGAGAAC-----
QY 225 AlaSerThrAsnAsnGlyGlnPheGlu 233
Db 556 -----AATGGGCGAAGATCCAG 573

RESULT 4
BE036639
LOCUS
DEFINITION
MP03B03 MP Mesembryanthemum crystallinum cDNA 5' similar to
ribosome-inactivating protein gelonin, mRNA sequence.

```

BE036639 1021 bp mRNA linear EST 07-JUN-2000  
 MP03B03 MP Mesembryanthemum crystallinum cDNA 5' similar to  
 ribosome-inactivating protein gelonin, mRNA sequence.

QY 161 ThrLeuPheTyr-----TyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeu 177  
... ||||| :||| |||||  
... ||||| :||| |||||  
Db 100 GTAGTTCCATTGAA---ACAGTCAGTTCGGTCTCACTGGGACCACCAGCAAAACCTAT 118

```

QY 38 GlyValPheIleSerAsnLeuArgLysAlaLeuProAsnGluArgLysLeuTyrAspIle 57
Db 157 AGTGCTTTCTCAGACCGCTTAAAGCAGTGAAGTAGTGAATGATAT 216
QY 58 ProLeuLeuArgSerSerLeuProGlySerGlnArgTyrAlaLeuIleHisLeuThrAsn 77
Db 217 CTTTGTCTGCTGGCAATCCGGTCCAGCAGGATTTGCTCTTGTGCAACTCTTTGAT 276
QY 78 TyrAlaAspGluThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyr 97
Db 277 TGGGACAAACACCGCATCAGCTGGTGCTTAATCGCGTCAATGCTTATCTGTAGCTTAT 336
QY 98 ArgAlaGlyAspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaLysTyr 117
Db 337 CAAGCTAAAAAATCGTTATTACTTGCTT-----TCGACACTCCCGCCAATCCTCAG 387
QY 118 ValPheLysAspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeu 137
Db 388 CTTTATGGAGCAACCCGACCGC-----CTTAGTTTACAGGCGAGCTACATTCGACTT 441
QY 138 GlnThrAlaAlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSer 157
Db 442 CAAAATGTTGCGAAACAGTAGAGAAAATATTGATCTGGGAATCAATCCACTTGCACG 501
QY 158 AlaIleThrLeuPheTyrTyrAsnAla-----AsnSerAlaAlaSerAla 173
Db 502 GCAATCAGACACTCCAACTCCGCGTCCGCTACGCTGGAGACATCCGCTGCTCGTTCC 560
QY 174 LeuMetValLeuIleGlnSerThrSerGluAlaAlaArgTyrLysPheIleGluGln 193
Db 561 CTTATCGTCTTATGATCAACTGCTCCGAGACCGCAGCATTCAGAGCGATCGACACGAG 620
QY 194 Ile 194
Db 621 TTA 623

RESULT 6
BQ583480
LOCUS E011979-024-005-N05-SP6 MP1Z-ADIS-024-inflorescence Beta vulgaris
DEFINITION cDNA clone 024-005-N05 5-PRIME, mRNA sequence.
ACCESSION BQ583480
VERSION BQ583480.1 GI:26113057
KEYWORDS EST.
SOURCE Beta vulgaris
ORGANISM Beta vulgaris
REFERENCE 1 (bases 1 to 658)
AUTHORS Herwig,R., Schulz,B., Weisshaar,B., Hennig,S., Steinfath,M.,
Drungowski,M., Stahl,D., Wruck,M., Menze,A., O'Brien,J., Lehrach,H.
and Radelof,U.
TITLE Construction of a 'unigene' cDNA clone set by oligonucleotide
fingerprinting allows access to 25 000 potential sugar beet genes
JOURNAL Plant J. 32 (5), 845-857 (2002)
MEDLINE 22362189
PUBMED 12472698
COMMENT Contact: Weisshaar B
ADIS DNA core facility at MP1Z
Max-Planck-Institute for Plant Breeding Research
Carl-von-Linne Weg 10, 50829 Koeln, Germany
Fax: 00492215062851
Email: weisshaar@mpiz-koeln.mpg.de
Insert Length: 658 Std Error: 0.00
Plate: 5 row: N column: 05
Seq primer: SP6; CATACGATTAGGTGACACTATAG.
Location/Qualifiers
1. .658
/organism="Beta vulgaris"
/mol type="mRNA"
/cultivar="KWS2320 (double haploid, monogerm breeding
line)"

```

```

/db_xref="GABI:183219"
/db_xref="taxon:161934"
/clone="024-005-N05"
/tissue_type="inflorescence"
/lab_host="EMDH10B"
/clone_lib="MP1Z-ADIS-024-inflorescence"
/note="Vector: pCMVSPORT6; Site 1: SalI; Site 2: NotI;
cDNA library from sugar beet, library provided by KWS
Kleinwanzlebener Saatgut AG Einbeck, Germany, contact:
b.schulz@kws.de; cloning sites SalI-NotI, primer sites and
orientation:
SP6-Sali-CCACGGCTCG-5prime-cDNA-polyA-CC-NotI-T7; Note:
Sequencing granted in the context of the GABI-Beet
project, local PI: Dr. Katharina Schneider, coordinator:
Prof. Christian Jung; Sequence submission managed by
RZPD/GABI-Primary database: http://gabi.rzpd.de"

ORIGIN
Alignment Scores:
Pred. No.: 4,56e-14 Length: 658
Score: 198.50 Matches: 57
Percent Similarity: 49.70% Conservative: 26
Best Local Similarity: 34.13% Mismatches: 75
Query Match: 13.86% Indels: 9
DB: 13 Gaps: 5

US-09-905-247-1 (1-289) x BQ583480 (1-658)

QY 26 SerPheArgLeuSerGlyAlaThrSerSerTyrGlyValPheIleSerAsnLeuArg 45
Db 166 ACATACGACCTGAGCGGGGTACA--AAAGCTACTCTCAATCTTGCAAAATTTAGC 222
QY 46 LysAlaLeuProAsnGluArgLysLeuTyrAspIleProLeuLeuArgSerSerLeuPro 65
Db 223 AAAAAATTG---GAAGCAACAAGGCTTGCAACATGCCAGTCGACGAGCAAACTTACA 279
QY 66 GlySerGlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAspGluThrIleSerVal 85
Db 280 GCAGCGCGGAATATGTTTGTAGTGATCTCAAGTTACCCAGTACAATATGTCACTTA 339
QY 86 AlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGlyAspThrSerTyr--- 104
Db 340 GCTATAGAGTTAATAATGTTAATGTTGCTGCTTATCGTGTGGCGGCAAACTCAAGC 399
QY 105 ---PhePheAsnGlyAlaSerAlaThrGluAlaLysTyrValPheLysAspAlaMet 123
Db 400 ACATTCTCTTATGAGCCAGATGTTACATTTGCCAGGCTTAAGCTTTTGTATAAGATGAT 459
QY 124 ArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAlaAlaGlyLys 143
Db 460 AATCAACCAACACTTCGATATGACAGCAACTACCGCTAATATAGAG--GCGGTGGTGGC 516
QY 144 IleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThrThrLeuPhe 163
Db 517 TCCAGGAAGGACCTTCTCTGGGCTTCAGAAATTACGAGAGAAATTTAGCTCTTT 576
QY 164 TyrTyrAsn-----AlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGln 179
Db 577 GACCAAAATAAACCAAGTGTGAGCGCGGAGGAGCTCAATTCTTATCAAGCATTTCAA 636
QY 180 SerThrSerGluAlaAlaArg 186
Db 637 ATGTTATCTGAGGCTGCGCA 657

RESULT 7
CAL198032
LOCUS SCEAD1078B05.g AD1 Saccharum officinarum cDNA clone SCEAD1078B05
DEFINITION S, mRNA sequence.
ACCESSION CAL198032
VERSION CAL198032.1 GI:35228882
KEYWORDS EST.
SOURCE Saccharum officinarum

```

ORGANISM Saccharum officinarum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Saccharum.

REFERENCE 1 (bases 1 to 601)  
Vettore, A.L., da Silva, F.R., Kemper, E.L. and Arruda, P.  
The libraries that made SUCEST  
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)  
Contact: Arruda P  
Centro de Biologia Molecular e Engenharia Genetica  
Universidade Estadual de Campinas  
Caixa Postal 6010, 13083-970, Campinas SP, Brazil  
Tel: 55 19 3788 1137  
Fax: 55 19 3788 1089  
Email: parruda@unicamp.br  
Clone distribution: Clone distribution information can be found through the Brazilian Clone Collection Center (BCCC) at <http://www.bcccnet.fcap.unesp.br>  
Plate: 078 row: B column: 05  
Seq primer: T7 Promoter Primer.  
Location/Qualifiers  
1..601  
/organism="Saccharum officinarum"  
/mol\_type="mRNA"  
/db\_xref="taxon:4547"  
/clone="SCEZAD1078B05"  
/lab\_host="DH10B"  
/clone\_lib="AD1"  
/note="Organ: seedlings inoculated with Gluconacetobacter diazotrophicans; Vector: pSport1; Site 1: SalI; Site 2: NotI; An unidirectional cDNA library generated from [seedlings inoculated with Gluconacetobacter diazotrophicans]. cDNA was prepared from polyA+ mRNA using SuperScript Plasmid System Kit (Invitrogen). The double-strand cDNAs were fractionated in a Sepharose CL-2B 40cm-column and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at <http://sucest.lad.ic.unicamp.br/public>

FEATURES  
source  
1..1033  
/organism="Mesembryanthemum crystallinum"  
/mol\_type="mRNA"  
/db\_xref="taxon:3544"  
/tissue\_type="apical meristem and leaf primordia"  
/dev\_stage="6 weeks"  
/clone\_lib="Mp"  
/note="3 d 500mM NaCl"

ORIGIN  
US-09-905-247-1 (1-289) x CAL198032 (1-601)  
QY 91 AsnValTyrIleMetGlyTyrArgAlaGlyAspThrSerTyrPheAsnGlyAlaSer 110  
Db 3 AATGTTTACATCATCGTTATCTTATGTTACACAGATATCTTCTCAGGAGCTGAA 62  
QY 111 AlaThrGluAlaAlaLysTyrValPheLysAspAlaMetArgLysValThrLeuProTyr 130  
Db 63 ---GATGGGAGATCGGCTTCTTCACAGATGACCCAGCCAGCTTCAGACACAGGAGA 119  
QY 131 SerGlyAsnTyrGluArgLeuGlnThrAlaAlaGlyLysIleArgGluAsnIleProLeu 150  
Db 120 TCAGCCAGCTATGCGCAGATGAGAGCCGAGGACACATAAATAGGGCTAACATCGACCTA 179  
QY 151 GlyLeuProAlaLeuAspSerAlaIleThrThrLeuPheTyrTyrAsnAlaAsnSerAla 170  
Db 180 GGTCTGCTAGTATGAGCGGCCCATAGAACCTTTAGCTGGAGACACAGAGATGTA 239  
QY 171 AlaSerAlaLeuMetValLeuIleGlnSerThrSerGluAlaAlaArgTyrLysPheIle 190  
Db 240 AGGTCA---CTTCTAGTTGTTATCCAAATGGTTGCAGAGGGTGTGGGTTTGGGAGATA 296  
QY 191 GluGlnGlnIleGlyLysArgValAspLysThrPheLeuProSerLeuAlaIleSer 210

297 GAGTCGTACACCTCCACAG-----TCTTCATTTCGACCTAGCCACGATGATTGGT 350  
211 LeuGluAsnSerTyrPheAlaLeuSerLysGlnIleAlaSerThrAsnAsnGly 230  
351 TTTGAGAACCACTGGGAGACCTCTCTGAACAGATCAAACTAGCGAGGTA-----TTC 404  
231 GlnPheGluSerProValValLeuIleAsnAlaGlnAsnGlnArgValThrIleThrAsn 250  
405 GCCTTCCAACTGCT-----ATTAACTTTGGCGGAA-----GTTGTGACAAT 449  
251 ValAspAlaGlyValThrSerAsnIleAlaLeuLeu 263  
450 GTTCGATCAGCAGTTATAAGGCTTTGTTCTTAATGCTG 488  
RESULT 8  
LOCUS BE036541 1033 bp mRNA linear EST 07-JUN-2000  
DEFINITION MP01807 MP Mesembryanthemum crystallinum cDNA 5' similar to antiviral protein, mRNA sequence.  
ACCESSION BE036541  
VERSION BE036541.1 GI:8331550  
KEYWORDS EST.  
SOURCE Mesembryanthemum crystallinum (common iceplant)  
ORGANISM Mesembryanthemum crystallinum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Caryophyllales; Aizoaceae; Mesembryanthemum.  
REFERENCE 1 (bases 1 to 1033)  
AUTHORS Bohnert, H.J., Borchert, C., Brazille, S., Brooks, J., Eaton, M., Ferreira, H., Kawasaki, S., McColough, A., Michalowski, C.B., Palacios, C., Scarra, G., Wheeler, M. and Zepeda, G.R.  
TITLE Functional Genomics of Plant Stress Tolerance  
JOURNAL Unpublished (2000)  
COMMENT Contact: Michalowski, C.B.  
University of Arizona  
Bio Sciences West room 513, Tucson, AZ 85721, USA  
Tel: 520-621-7982  
Fax: 520-621-1697  
Email: chm@u.arizona.edu  
An open reading frame exists.  
Location/Qualifiers  
1..1033  
/organism="Mesembryanthemum crystallinum"  
/mol\_type="mRNA"  
/db\_xref="taxon:3544"  
/tissue\_type="apical meristem and leaf primordia"  
/dev\_stage="6 weeks"  
/clone\_lib="Mp"  
/note="3 d 500mM NaCl"

ORIGIN  
US-09-905-247-1 (1-289) x BE036541 (1-1033)  
QY 98 AspValThrAsnValTyrIleMetGlyTyrArg-----AlaGlyAspThrSer 103  
Db 38 GAGGCGCAACAGGCTTATGTCGTGCTACCGCAAGCTTGGCGGAAAGACCGTGCC 97  
QY 104 TyrPhePheAsnGlyValSerAlaThrGluAlaAlaLysTyrValPheLysAspAlaMet 123  
Db 98 AACTTCCTTAGTGATGCT---CCACGGTGGCTAGAACCACTCTTCAAGGCGGAGC 154  
QY 124 ArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAlaAlaGlyLys 143  
Db 155 GTTCGG---AACATTGCGTTCGAGGAGGTTACATAGCTTAGAAGAGCTGCAAGCAA 211





/cell\_type="epidermal bladder cells"  
 /dev\_stage="12 weeks old"  
 /clone\_lib="WM"  
 /notes="Vector: Bluescript SK+; Site\_1: EcoRI; Site\_2:  
 XhoI; Plants stressed 6 weeks in 500mM NaCl"

## ORIGIN

Alignment Scores:  
 Pred. No.: 1.45e-11 Length: 993  
 Score: 181.00 Matches: 55  
 Percent Similarity: 51.76% Conservative: 33  
 Best Local Similarity: 32.35% Mismatches: 66  
 Query Match: 12.64% Indels: 16  
 DB: 10 Gaps: 7

US-09-905-247-1 (1-289) x BE0335039 (1-993)

QY 101 AspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPheLys 120  
 DB 787 GATCGAGCCAACTTCCTAGGATGCT--CCTACAGTCGCTAAGAAACCTTCTTTAAG 731  
 QY 121 AspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAla 140  
 DB 730 GGAGCGACCGTTCGA--AACATTCATTTGGTGGTAACATACATAGCTCTGGAGAAAGCT 674  
 QY 141 AlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThr 160  
 DB 673 GCCGGTCAAGACGAAACAGCATATTAGTTGGGGCTTGTAACTAGATTGGCATCGAG 614  
 QY 161 ThrLeuPhe-----TyrTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeu 177  
 DB 613 TCGATTATGTATAAAACGATCGATGGAAACCTCGAGGCCAAATCTTACTGATTGCC 554  
 QY 178 IleGlnSerThrSerGluAlaAlaArgTyrLysPheIleGluGlnGlnIleGlyLysArg 197  
 DB 553 ATCCAGATGGTTTCAGAGACGACGGTTCATATATCGAGACTAAGGTGACCGAAAGT 494  
 QY 198 ValAspLys---ThrPheLeuProSerLeuAlaIleIleSerLeuGluAsnSerTyrSer 216  
 DB 493 GGGTCACATGGTTGTTCAAAACCGATCGGAAAGTATTAATCTGGAGAACCACTGGGGA 434  
 QY 217 AlaLeuSerLysGlnIleGlnIleAlaSerThrAsn-----AsnGlyGlnPhe 232  
 DB 433 AAGATTTCGGATGAGATTTCATAGTCAGTTCAGTTCGAAACCACTGCTAAATGTACGAAC 374  
 QY 233 GluSerProValVal---LeuIleAsnAlaGlnAsn-----GlnArgVal 246  
 DB 373 ATTTCCCTCGGATTAATCAATCTGATGCTGATGCTACACATGGAAGTGGATAAGATT 314  
 QY 247 ThrIleThrAsnValAspAlaGlyValVal 256  
 DB 313 GCCACTATAAACCTGACTTGGGGATCTC 284

BE033546 1038 bp mRNA linear EST 07-JUN-2000  
 MF03A09 MF Mesembryanthemum crystallinum cDNA 5' similar to  
 ribosome-inactivating protein, mRNA sequence.

ACCESSION BE033546

VERSION BE033546

KEYWORDS EST.

SOURCE BE033546.1 GI:8328555

ORGANISM Mesembryanthemum crystallinum (common iceplant)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Caryophyllales; Alzooceae; Mesembryanthemum.

REFERENCE 1 (bases 1 to 1038)

AUTHORS Bohnert, H.J., Borchert, C., Brazille, S., Brooks, J., Eaton, M.,

Ferreira, H., Kawasaki, S., McColough, A., Michalowski, C.B.,

Palcio, C., Scara, G., Wheeler, M. and Zepeda, G.R.

Functional Genomics of Plant Stress Tolerance

Unpublished (2000)

CONTACT: Michalowski, C.B.

University of Arizona  
 Bio Sciences West room 513, Tucson, AZ 85721, USA  
 Tel: 520-621-7982  
 Fax: 520-621-1697  
 Email: cbm@u.arizona.edu  
 An open reading frame exists.

## FEATURES

Location/Qualifiers  
 1..1038  
 /organism="Mesembryanthemum crystallinum"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3544"  
 /tissue\_type="Root"  
 /dev\_stage="5-6 weeks old"  
 /clone\_lib="MF"  
 /notes="Vector: Bluescript SK+; Site\_1: EcoRI; Site\_2:  
 XhoI"

## ORIGIN

Alignment Scores:  
 Pred. No.: 4.33e-11 Length: 1038  
 Score: 177.50 Matches: 65  
 Percent Similarity: 49.50% Conservative: 34  
 Best Local Similarity: 32.50% Mismatches: 74  
 Query Match: 12.40% Indels: 27  
 DB: 10 Gaps: 8

US-09-905-247-1 (1-289) x BE033546 (1-1038)

QY 56 AspIleProLeuLeuArgSerSerLeuProGly-SerGlnArgTyrAlaLeuIleHisLe 75  
 DB 33 GATCNCGGCTCGCAGGAATTCGGTCGGTCGCGACCAACGATTCCTCTCGTCGACCT 92  
 QY 75 uThrAsnTyrAlaAspGluThrIleSerValAlaIleAspValThr-AsnValTyrIleM 95  
 DB 93 CAHAACAACCTCGCAAGACCATCACACTCGCAATCGCGTCGACAGAGCTCTATGTCTG 152  
 QY 95 etGlyTyrArg-----AlaGlyAspThrSerTyrPhePheAsnGlyAlaSerA 111  
 DB 153 TGGCTACTCCGCGACAAGCTTCGCGAAAAGACCTGCACTTCTTACGATGCT---C 209  
 QY 111 laThrGluAlaAlaLysTyrValPheLysAspAlaMetArgLysValThrLeuProTyrS 131  
 DB 210 CCACGTCGCTAGGAACAACTCTTCAAGGCGCGACGGTTCGG---AACATGGGTCG 266  
 QY 131 erGlyAsnTyrGluArgLeuGlnThrAlaAlaGlyLysIleArgGluAsnIleProLeuG 151  
 DB 267 GAGGAGTTACATAGCTTAGAGAGGCTGCAAGCAAGCGCGAAATGCAATAGAGTTAG 326  
 QY 151 lyLeuProAlaLeuAspSerAlaIleThrThrLeuPhe-----TyrTyrAsnAlaA 168  
 DB 327 GGGTGAATAAACTAGAGTTCGCGATCGGCTTTTGGTAAGAACCCGATTATATGGCG 386  
 QY 168 snSerAlaAlaSerAlaLeuMetValLeuIleGlnSerThrSerGluAlaAlaArgTyrL 188  
 DB 387 AGGTGAGCCCAAAATCTTCTGCTGATGCTATACAGATGTTCTTCTGAAGCAGCACGTTTA 446  
 QY 188 yspPheIleGluGlnIleGlyLysArg---ValAspLysThrPheLeuProSerLeuA 207  
 DB 447 AGTTTATTGAAAGTAAAGTACCACCAAGTGGTGTATCATGCTGTTGTTCAACCCGCGA 506  
 QY 207 laIleIleSerLeuGluAsnSerTyr-----SerAlaA 218  
 DB 507 AAGTCTGAGTTCGGTGTACATTCGGTGAACATTTTCATGAGACTCATAGACAGCGG 566  
 QY 218 euSerLysGlnIleGlnIleAlaSerThrAsnAsnGlyGlnPheGluSerProVal 236  
 DB 567 CGTGAACACCT-----GCCTGACTAAT-----TTCCACGCGCTATT 604

RESULT 12

CA838926

LOCUS

DEFINITION

CA838926 698 bp mRNA linear EST 12-DEC-2002  
 MCT021H09 170947 Ice plant Lambda Uni-zap XR expression library, 5  
 days 0.5 M NaCl treatment, Crassulacean acid metabolism, phase IV

**OY** 99 -----AlaGlyAspThrSerTyrPhePheAsnGlyAlaSerAlaThrGlu 113  
||| :| | ||| ||| |||  
...CGAGCAGACGGGCTATGTCGCCCTATCCGCACAAGTTCAAA 433

```

/mol_type="mRNA"
/db_xref="taxon:3544"
/clone="MCT036G03"
/tissue_type="leaf"
/dev_stage="five-week-old"
/clone_lib="Ice plant Lambda Uni-Zap XR expression
library, 5 days 0.5 M NaCl treatment, Crassulacean acid
metabolism, phase IV (5:30 PM)"
/notes="Vector: Lambda Uni-Zap XR, Bluescript SK-; Site 1:
EcoRI; Site 2: XhoI; Library construction was performed
according to Stratagene's recommended protocol for the
Lambda UniZapXR vector and cDNA synthesis kit."

/mot_type="mRNA"
/db_xref="taxon:3544"
/clone="MCT036G03"
/tissue_type="leaf"
/dev_stage="five-week-old"
/clone_lib="Ice plant Lambda Uni-Zap XR expression
library, 5 days 0.5 M NaCl treatment, Crassulacean acid
metabolism, phase IV (5:30 PM)"
/notes="Vector: Lambda Uni-Zap XR, Bluescript SK-; Site 1:
EcoRI; Site 2: XhoI; Library construction was performed
according to Stratagene's recommended protocol for the
Lambda UniZapXR vector and cDNA synthesis kit."

US-09-905-247-1 (1-289) x CA838446 (1-667)
QY 25 ValSerPheArgLeuSerGlyAlaThrSerSerTyrGlyValPheIleSerAsnLeu 44
Db 209 GTGACATTCGACCTC-----GCCCAATCCGACACATATTCGAAGTTTCATTCGCCGATCTC 262
QY 45 ArgLysAlaLeuProAsnGluArgLysLeuTyrAspIleProLeuLeuArgSerSerLeu 64
Db 263 CAAACTGAACTATTCTCCGGCCACACCGCATGTCAAATCCCGTGCACGCGGACACCGCA 322
QY 65 ProGlySer-----GlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAspGluThr 82
Db 323 ACCACATCAATGACCAACGATTCGCTCGTCCGACCTCGAAACACCTCGCAAAAACC 382
QY 83 IleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArg----- 98
Db 383 ATCACCCTCGCAATGACGCTGACAGACGCTACGTCGCTCGCTATCGGACAGAGTTCAA 442
QY 99 -----AlaGlyAspThrSerTyrPhePheAsnGlyValAsnAlaThrGlu 113
Db 443 GAAAGACCGCAACATGAAAGGACGTCGCAACTTCCTTCAGATGTCATCT---ACGGTT 499
QY 114 AlaAlaLysTyrValPheLysAspAlaMetArgLysValThrLeuProTyrSerGlyAsn 133
Db 500 GCGAAGGAGAACTCTTTTCATGAGCAACAGTTTCGG---AACCTTCGCTTCAAAGGGACT 556
QY 134 TyrGluArgLeuGlnThrAlaAlaGlyLysIleArgGluAsnIleProLeuGlyLeuPro 153
Db 557 TACACATCTTTAGAGATGCTGCAAAACCAACGACGAGAGCCATAGATGATTAGGGTGGAT 616
QY 154 AlaLeuAspSerAlaIleThrThrLeuPhe 163
Db 617 AAACATAGATTGCGATCGAGTCGGTTAT 646

RESULT 14
CA840373
LOCUS
DEFINITION
MCT036G03 173841 Ice plant Lambda Uni-Zap XR expression library, 5
days 0.5 M NaCl treatment, Crassulacean acid metabolism, phase IV
(5:30 PM). Mesembryanthemum crystallinum cDNA clone MCT036G03 5,
mRNA sequence.
CA840373.1 GI:26568138
ACCESSION
VERSION
KEYWORDS
SOURCE
Mesembryanthemum crystallinum (common iceplant)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Caryophyllales; Aizoaceae; Mesembryanthemum.
1 (bases 1 to 678)
Cushman, J.C.
An expressed sequence tag database for the common ice plant,
Mesembryanthemum crystallinum
Unpublished (1997)
Contact: Cushman JC
Department of Biochemistry
University of Nevada
MS200, Reno, NV 89557-0014, USA
Tel: 775-784-1918
Fax: 775-784-1650
Email: jcushman@unr.edu
PCR Primers
FORWARD: T3 20mer
BACKWARD: T7 21mer
Plate: 036 row: G column: 03
Seq primer: T3 20mer
High quality sequence stop: 678.
Location/Qualifiers
1..678

/mot_type="mRNA"
/db_xref="taxon:3544"
/clone="MCT036G03"
/tissue_type="leaf"
/dev_stage="five-week-old"
/clone_lib="Ice plant Lambda Uni-Zap XR expression
library, 5 days 0.5 M NaCl treatment, Crassulacean acid
metabolism, phase IV (5:30 PM)"
/notes="Vector: Lambda Uni-Zap XR, Bluescript SK-; Site 1:
EcoRI; Site 2: XhoI; Library construction was performed
according to Stratagene's recommended protocol for the
Lambda UniZapXR vector and cDNA synthesis kit."

/mot_type="mRNA"
/db_xref="taxon:3544"
/clone="MCT036G03"
/tissue_type="leaf"
/dev_stage="five-week-old"
/clone_lib="Ice plant Lambda Uni-Zap XR expression
library, 5 days 0.5 M NaCl treatment, Crassulacean acid
metabolism, phase IV (5:30 PM)"
/notes="Vector: Lambda Uni-Zap XR, Bluescript SK-; Site 1:
EcoRI; Site 2: XhoI; Library construction was performed
according to Stratagene's recommended protocol for the
Lambda UniZapXR vector and cDNA synthesis kit."

US-09-905-247-1 (1-289) x CA840373 (1-678)
QY 25 ValSerPheArgLeuSerGlyAlaThrSerSerTyrGlyValPheIleSerAsnLeu 44
Db 211 GTGACATTCGACCTC-----GCCCAATCCGACACATATTCGAAGTTTCATTCGCCGATCTC 264
QY 45 ArgLysAlaLeuProAsnGluArgLysLeuTyrAspIleProLeuLeuArgSerSerLeu 64
Db 265 CAAACTGAACTATTCTCCGGCCACACCGCATGTCAAATCCCGTGCACGCGGACACCGCA 324
QY 65 ProGlySer-----GlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAspGluThr 82
Db 325 ACCACATCCCAATGACCAACGATTCGCTCGTCCGACCTCGAAACACCTCGCAAAAACC 384
QY 83 IleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArg----- 98
Db 385 ATCACCCTCGCAATCGACGTCAGACAGCTCTAGCTGCTCGCTTCGCGCATCGCAAGATTCAA 444
QY 99 -----AlaGlyAspThrSerTyrPhePheAsnGlyAlaSerAlaThrGlu 113
Db 445 GAAAGACCGGACCAAAATGAAAGGACCGTCGCAACTTCCTTCGAATGCACTCT---ACGGTT 501
QY 114 AlaAlaLysTyrValPheLysAspAlaMetArgLysValThrLeuProTyrSerGlyAsn 133
Db 502 GCGAAGGAGAACTCTTTTCATGAGCAACAGTTTCGG---AACCTTCGCTTCAAAGGGACT 558
QY 134 TyrGluArgLeuGlnThrAlaAlaGlyLysIleArgGluAsnIleProLeuGlyLeuPro 153
Db 559 TACACATCTTTAGAGATGCTGCAAAACCAACGACGAGAGCCATAGATGATTAGGGTGGAT 618
QY 154 AlaLeuAspSerAlaIleThrThrLeuPhe 163
Db 619 AAACATAGATTGCGATCGAGTCGGTTAT 648

RESULT 15
BE034055
LOCUS
DEFINITION
MG04C05 MG Mesembryanthemum crystallinum cDNA 5' similar to
Ribosome-inactivating protein gelonin, mRNA sequence.
BE034055
ACCESSION
VERSION
KEYWORDS
SOURCE
Mesembryanthemum crystallinum (common iceplant)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Caryophyllales; Aizoaceae; Mesembryanthemum.
1 (bases 1 to 811)
Bohnert, H.J., Borchert, C., Brazille, S., Brooks, J.J., Eaton, M.,
Perrea, H., Kawasak, S., McColough, A., Michalowski, C.B.,
Palacio, C., Scara, G., Wheeler, M. and Zepeda, G.R.

```

TITLE Functional Genomics of Plant Stress Tolerance  
JOURNAL Unpublished (2000)  
COMMENT Contact: Michalowski, C.B.  
University of Arizona  
Bio Sciences West room 513, Tucson, AZ 85721, USA  
Tel: 520-621-7982  
Fax: 520-621-1697  
Email: cbm@u.arizona.edu  
An open reading frame exists.  
Insert length: 1 Std Error: 0.00.

## FEATURES

source  
1..811  
/organism="Mesembryanthemum crystallinum"  
/mol\_type="mRNA"  
/db\_xref="taxon:3544"  
/tissue\_type="roots"  
/dev\_stage="5-6 weeks"  
/clone\_lib="MG"  
/note="grown in hydroponics, stress 400 mM NaCl (in 0.5  
Hoagland's), 30 h stress"

## ORIGIN

## Alignment Scores:

Pred. No.:	2 97e-10	Length:	811
Score:	169.50	Matches:	59
Percent Similarity:	50.81%	Conservative:	35
Best Local Similarity:	31.89%	Mismatches:	69
Query Match:	11.84%	Indels:	22
DB:	10	Gaps:	7

US-09-905-247-1 (1-289) x BE034055 (1-811)

QY	25	ValSerPheArgLeuSerGlyAlaThrSerSerSerTyrGlyValPheIleSerAsnLeu	44
Db	252	GTGACATTCAACCTC-----GCCCAACCCACACATATTCACATTCATGACATCTCTC	305
QY	45	ArgLysAlaLeuProAsnGluArgLysLeuTyrAspIleProLeuLeuArgSerSerLeu	64
Db	306	CGAGCTCAACTCTGTGGCAACACCGCATGT---CAAATCCCGTGACACGGTCAACCGCA	362
QY	65	ProGlySerGlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAspGluThrIleSer	84
Db	363	CCGGATCTCCACGATTCCTCTCGTCGACCTCAAAACCAACCTCGCAAAAGACCATCACA	422
QY	85	ValAlaIleAspValThrAsnValTyrIleMetGlyTyrArg-----AlaGly	100
Db	423	CTCGCAATCGAGTGCACAGCGTCTATGTCGTGGCTACCGGACACAGCTTGGCGAATA	482
QY	101	AspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPheLys	120
Db	483	GACCGTGCCAACTTCTTAGCGATGCT---CCACGGTCGTAGGAACAACCTTTCAG	539
QY	121	AspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAla	140
Db	540	GGCGGACGGTTCGG---AACATTCCGTCGGAGGGAGTTACATAGCCCTAGAAAAAGCT	596
QY	141	AlaGlyLysIleArgGluAsnIle-ProLeuGlyLeuProAlaLeuAspSerAlaIleLeu	160
Db	597	GCAAGCAAGCGCGAATGCAATATGAGTTAGGTAGGGTGAATTAACTAGATTGCGATCGA	656
QY	160	rThrLeuPheTyrTyrAsn-----AlaAsnSerAlaAlase	172
Db	657	ATCCGTTTTTGGGAAGAACCCCATTTATGTGTCAGGTTGAAGGCCAAATCTTGCTGATTG	716
QY	172	rAlaLeuMetValLeuIleGlnSerThrSerGluAlaAlaArgTyrLysPheIleGluG1	192
Db	717	CTATACCAAGATTGGTCTC-----GAGCAACACGGTTTTTAGTTTTTGAAG	764
QY	192	nGlnIleGlyLys	196
Db	765	TACGGTACCCAAA	777

Search completed: September 14, 2004, 21:56:37  
Job time : 2484 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 14, 2004, 21:15:17 ; Search time 455 Seconds

(without alignments)  
3195.846 Million cell updates/sec

Title: US-09-905-247-1

Perfect score: 1432

Sequence: 1 MIRFLVLSLLILTLFLTPA.....AAMDDVPMVTFGCGSYAL 289

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 3304383 seqs, 2515761380 residues

Total number of hits satisfying chosen parameters: 6608766

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool\_p/US0905247/rnat\_10092004\_103936\_4829/app\_query.fasta\_1.455  
-DB=Published Applications NA -QEMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1  
-LOOPECL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=Blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0  
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DRV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-XGAPOP=6 -XGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA.\*

1:	/cgn2_6/ptodata/1/pubna/US07_PUBCOMB.seq.*
2:	/cgn2_6/ptodata/1/pubna/PCT_NEW_PUB.seq.*
3:	/cgn2_6/ptodata/1/pubna/US06_NEW_PUB.seq.*
4:	/cgn2_6/ptodata/1/pubna/US06_PUBCOMB.seq.*
5:	/cgn2_6/ptodata/1/pubna/US07_NEW_PUB.seq.*
6:	/cgn2_6/ptodata/1/pubna/PCTUS_PUBCOMB.seq.*
7:	/cgn2_6/ptodata/1/pubna/US08_NEW_PUB.seq.*
8:	/cgn2_6/ptodata/1/pubna/US08_PUBCOMB.seq.*
9:	/cgn2_6/ptodata/1/pubna/US09A_PUBCOMB.seq.*
10:	/cgn2_6/ptodata/1/pubna/US09B_PUBCOMB.seq.*
11:	/cgn2_6/ptodata/1/pubna/US09C_PUBCOMB.seq.*
12:	/cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq.*
13:	/cgn2_6/ptodata/1/pubna/US09_PUB.seq.*
14:	/cgn2_6/ptodata/1/pubna/US10A_PUBCOMB.seq.*
15:	/cgn2_6/ptodata/1/pubna/US10B_PUBCOMB.seq.*
16:	/cgn2_6/ptodata/1/pubna/US10C_PUBCOMB.seq.*
17:	/cgn2_6/ptodata/1/pubna/US10_PUB.seq.*
18:	/cgn2_6/ptodata/1/pubna/US60_NEW_PUB.seq.*
19:	/cgn2_6/ptodata/1/pubna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	ID	Description
------------	-------	--------------------	----	-------------

#### ALIGNMENTS

RESULT 1  
US-10-280-725B-3  
; Sequence 3, Appli US-10-280-725B-3  
; Publication No. US20040049025A1  
; GENERAL INFORMATION:  
; APPLICANT: Large Scale Biology Corporation  
; TITLE OF INVENTION: Recombinant Viral Nucleic Acids  
; FILE REFERENCE: LSBC-0109-US02  
; CURRENT APPLICATION NUMBER: US/10/280,725B  
; PRIOR FILING DATE: 2002-10-24  
; PRIOR FILING DATE: 09/557,941  
; PRIOR FILING DATE: 2000-04-24  
; PRIOR FILING DATE: 08/484,341  
; PRIOR FILING DATE: 1995-06-07  
; PRIOR APPLICATION NUMBER: 07/923,692  
; PRIOR FILING DATE: 1992-07-31  
; PRIOR APPLICATION NUMBER: 07/600,244  
; PRIOR FILING DATE: 1990-10-22  
; PRIOR APPLICATION NUMBER: 07/641,617  
; PRIOR FILING DATE: 1991-01-16  
; PRIOR APPLICATION NUMBER: 07/737,899  
; PRIOR FILING DATE: 1991-07-26  
; PRIOR APPLICATION NUMBER: 07/739,143





QY 21 ValGluGlyAspValSerPheArgLeuSerGlyAlaThrSerSerSerTyrGlyValPhe 40  
DB 68 GTGGAGGGCGATGTAGCTTCCTGTTATCAGGTGCAACAGCAGTTCCTATGAGTTTC 127  
QY 41 IleSerAsnLeuArgLysAlaLeuProAsnGluArgLysLeuTyrAspIleProLeuLeu 60  
DB 128 ATTTCAAACTGAGAAAAGCTCTTCCAAATGAAAGGAACTGACGATATCCCTCTGTGA 187  
QY 61 ArgSerSerLeuProGlySerGlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAsp 80  
DB 188 CGTTCCTCTCTCCAGTCTCTCAACGCTAGCGATTCATCTCAAAATTCAGCGCAT 247  
QY 81 GluThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGly 100  
DB 248 GAAACCATTTTCAGTGCCATAGACGTAACGAACTCTATATATGAGATATCGCTGCG 307  
QY 101 AspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPheLys 120  
DB 308 GATACATCTCTATTTTTCACGAGGCTCTGCAACAGAGCTGCAAAATATGATTTCAAA 367  
QY 121 AspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAla 140  
DB 368 GAGGCTATCGGAAAGCTTACGCTTCCATATCTGGCAATTCGAAAGGCTTCAAACTGCT 427  
QY 141 AlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThr 160  
DB 428 GCGGGCAAAATAAGGAAATATTCGCTTGGACTCCAGCTTTGGACAGTGGCATACC 487  
QY 161 ThrLeuPheTyrTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSer 180  
DB 488 ACTTGTCTTACTACACGCCAAATCTGCTGCTGGCACTTATGCTACTTCAATTCAGTCG 547  
QY 181 ThrSerGluAlaAlaArgTyrLysPheIleGluGlnGlnIleGlyLysArgValAspLys 200  
DB 548 AGCTCTGAGGCTCGGAGGTATTAATTTATTTAGCAACAATTTGGGAAGCGCTTGACAAA 607  
QY 201 ThrPheLeuProSerLeuAlaIleLeuSerLeuGluAsnSerTrpSerAlaLeuSerLys 220  
DB 608 ACCTTCTCTACCAAGTTAGCAATTAAGTTTGGAAATAGTTGCTGCTCTCTCCAG 567  
QY 221 GlnIleGlnIleAlaSerThrAsnGlnGlnPheGluSerProValValLeuIleAsn 240  
DB 668 CAATTCAGATAGCGAGTACTAATATGACAGAGTTTGAACCTCTGTTGTGCTTAAAT 727  
QY 241 AlaGlnAsnGlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnIle 260  
DB 728 GCTCAAAACCAACGAGTCATGATAACCAATGTTGATGCTGGAGTTGTAACTCCCAACATC 787  
QY 261 AlaLeuLeuLeuAsnArgAsnAsnMetAlaAlaMetAspAspValProMetThrGln 280  
DB 788 CGGTTGCTGCTGAATCGAAACAATATGCGAGCATGGATGACGATGTTCCCTATGACACAG 847  
QY 281 SerPheGlyCysGlySerTyrAlaLeu 289  
DB 848 AGCTTGGATGTGAAGTTATGCTATT 874

## RESULT 3

US-10-127-890-13

Sequence 13, Application US/10127890

Publication No. US20030166196A1

GENERAL INFORMATION:

APPLICANT: Better, Marc D.

Carroll, Stephen F.

Studnicka, Gary M.

TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating

NUMBER OF SEQUENCES: 173

CORRESPONDENCE ADDRESS:

ADDRESS: McAndrews, Held &amp; Malloy, Ltd.

STREET: 500 West Madison Street, 34th floor

CITY: Chicago

STATE: Illinois

COUNTRY: USA  
ZIP: 60661  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/127,890  
FILING DATE: 23-Apr-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/646,360  
FILING DATE: 13-MAY-1996  
APPLICATION NUMBER: PCT/US94/05348  
FILING DATE: 12-MAY-1994  
APPLICATION NUMBER: US 08/064,691  
FILING DATE: 12-MAY-1993  
APPLICATION NUMBER: US 07/988,430  
FILING DATE: 09-DEC-1992  
APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-JUN-1992  
APPLICATION NUMBER: US 07/787,567  
FILING DATE: 04-NOV-1991

ATTORNEY/AGENT INFORMATION:

NAME: McNicholas, Janet M.

REGISTRATION NUMBER: 32,918

REFERENCE/DOCKET NUMBER: 200-70.P4

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312/707-8889

TELEFAX: 312/707-9155

TELEX: 650 388-1248

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:

LENGTH: 913 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

SEQUENCE DESCRIPTION: SEQ ID NO: 13:

US-10-127-890-13

## Alignment Scores:

Pred. No.: 5,558-86 Length: 913

Score: 728.00 Matches: 156

Percent Similarity: 72.83% Conservative: 37

Best Local Similarity: 58.87% Mismatches: 70

Query Match: 50.94% Indels: 3

DB: 15 Gaps: 2

US-09-905-247-1 (1-289) x US-10-127-890-13 (1-913)

QY 1 MetIleArgPheLeuValLeuSerLeuLeuIleLeuThrLeuPheLeuThrProAla 20  
DB 10 ATGGTGAATTCCTTACTTCTTTTATTAATTCGCATCTTCATTTGTTCTCTACT 69  
QY 21 ValGluGlyAspValSerPheArgLeuSerGlyAlaThrSerSerSerTyrGlyValPhe 40  
DB 70 GCCAAGGCGATGTTAACTTCGATTTGTCGACTGCCACTGCAAAACCTCACAAAATTT 129  
QY 41 IleSerAsnLeuArgLysAlaLeuProAsnGluArgLysLeuTyrAspIleProLeuLeu 60  
DB 130 ATCGAAGATTCAGGCGCATCTTCCATTTAGCCATTAAGTGTATGATATACCTCTACTG 189  
QY 61 ArgSerSerLeuProGlySerGlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAsp 80  
DB 190 TATTCCTACTATTCGAGCTCCAGAGCTTCTATCTCTCGATCTTCAAGTTATGCATAT 249  
QY 81 GluThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGly 100  
DB 250 GAAACCATCTCGGTGGCCATAGATGTACCAACGTTTATGTTGTGGCGTATCGCACCCGC 309  
QY 101 AspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPheLys 120

Wed Sep 15 10:32:02 2004

Db 310 GATGATCCACCTTTTAA--GAATCTCTCTGAGCTTATTAACATCTTCAAA 366  
Qy 121 AspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAla 140  
Db 367 --GGTACGCGGAAATACACTGCCATATACCGGTAATATGAAATCTTCAACTGCT 423  
Qy 141 AlaGlyLysLeuArgGluAsnLeuProLeuGlyLeuProAlaLeuAspSerAlaLeuThr 160  
Db 424 GCACACAAAATAAGAGAGAAATATTGATCTTGGACTCCCTGCTTGGATGTGCGCATTTACC 483  
Qy 161 ThrLeuPheTyrTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuLeuGlnSer 180  
Db 484 ACATGTTTTATACATGCCCAATCTGCTCTCTGCAATGCTTGTACTAATCCAGAG 543  
Qy 181 ThrSerGluAlaAlaArgTyrLysPheLeuGlnGlnLeuGlyLysArgValAspLys 200  
Db 544 ACTGCAGAGCTGCAAGATTTAAGTATATCAGGACACGTTGCTAAGTATGTTGCCACT 603  
Qy 201 ThrPheLeuProSerLeuAlaLeuLeuSerLeuGluAsnSerTyrSerAlaLeuSerLys 220  
Db 604 AACTTTAAGCCCAATCTAGCCATCATAGCTTGGAAATCAATGCTCTGCTCTCTCCAA- 662  
Qy 221 GlnLeuGlnLeuAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuLeuAsn 240  
Db 663 CAATCTTTTGGCGCAGAAATCAAGGAGGAAATTTAGAAATCTCTGCACTTATAAA 722  
Qy 241 AlaGlnAsnGlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnLeu 260  
Db 723 CCTACCGGGAAACGGTTTCAAGTAACCAATGTTGATTCAGATGTTGTAAGGTAATATC 782  
Qy 261 AlaLeuLeuLeuAsn 265  
Db 783 AAATCTCTGCTGAAC 797

RESULT 4  
US-10-394-511-31  
; Sequence 31, Application US/10394511  
; Publication No. US20040009551A1  
; GENERAL INFORMATION:  
; APPLICANT: Borgford, Thor  
; TITLE OF INVENTION: Ricin-Like Toxin Variants for Treatment of Cancer,  
; TITLE OF INVENTION: Viral or Parasitic Infections  
; FILE REFERENCE: 10447-27  
; CURRENT APPLICATION NUMBER: US/10/394,511  
; PRIOR FILING DATE: 2003-03-24  
; PRIOR APPLICATION NUMBER: US 09/403,752  
; PRIOR FILING DATE: 1999-10-29  
; PRIOR APPLICATION NUMBER: US 60/063,715  
; PRIOR FILING DATE: 1997-10-29  
; PRIOR APPLICATION NUMBER: US 60/045,148  
; PRIOR FILING DATE: 1997-10-30  
; NUMBER OF SEQ ID NOS: 142  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 31  
; LENGTH: 1855  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: pAP-242 insert  
US-10-394-511-31

Alignment Scores:  
Pred. No.: 1.64e-45 Length: 1855  
Score: 427.00 Matches: 97  
Conservative: 58.76% Conservative: 64  
Best Local Similarity: 35.40% Mismatches: 93  
Query Match: 29.82% Indels: 20  
DB: 16 Gaps: 8

US-09-905-247-1 (1-289) x US-10-394-511-31 (1-1855)

Qy 25 ValSerPheArgLeuSerGlyAlaThrSerSerTyrGlyValPheLeuSerAsnLeu 44

Db 136 ATAAACTTTTACCACGCGGTGCCACTGTGCAAAAGCTACACAACCTTTTATCAGAGCTGTT 195  
Qy 45 ArgLysAlaLeuProAsnGluArgLysLeu--TyrAspLeuProLeuLeu--Arg 61  
Db 196 CGCGTCTTTTAACTGAGCTGATGTGAGACATGATATCAGTGTGTCACAAACAGA 255  
Qy 62 SerSerLeuProGlySerGlnArgTyrAlaLeuLeuHleHisLeuThrAsnTyrAlaAspGlu 81  
Db 256 GTTGTGTTGCTTATAAACCAACGCTTTATTATTAGTTGAACCTCTCAAAATCATGAGAGCTT 315  
Qy 82 ThrIleSerValAlaLeuAspValThrAsnValTyrIleMetGlyTyrArgAlaGlyAsp 101  
Db 316 TCTGTATCATAGCGCTGGATGTACCAATGCATATGTGTCGCTACCGTCTGCGAAAT 375  
Qy 102 ThrSerTyrPhePhe-----AsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPhe 119  
Db 376 AGCCATATTTCTTTCATCTCTGCAATCAGGAAGATCGAGAAGCAATCCTCATCTTTTC 435  
Qy 120 LysAspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThr 139  
Db 436 ACTGATGTTCAAAATCGATATACATTCGCCCTTTTGGTGGTAAATATGATAGACTTTGAA 495  
Qy 140 AlaAlaGlyLysLeuArgGluAsnLeuProLeuProAlaLeuAspSerAlaLeu 159  
Db 496 CTTGCTGTGTAATCTGAGAGAAATATCGAGTTGGAAATGTGTCCTCCTAGAGAGGCTATC 555  
Qy 160 ThrThrLeuPheTyrTyrAsnAlaAsn-----SerAlaAlaSerAlaLeu 174  
Db 556 TCAGCGCTTTTATTATACAGTACTGGTGGCACTCAGCTTCCAACTCTGCTCGTCTCTTT 615  
Qy 175 MetValLeuLeuGlnSerThrSerGluAlaAlaArgTyrLysPheLeuGlnGlnLeu 194  
Db 616 ATAAATTTGCATCCAAATGATTTTCAGAGCAGCAAGATTTCCAAATATATTGAGGGGAAATG 675  
Qy 195 GlyLysArgVal-----AspLysThrPheLeuProSerLeuAlaLeuLeuSerLeuGlu 212  
Db 676 CGCAGGAATTAGGTACAAACCGGAGATCTGCACAGATCTTAGCGTAAATACACTTGAG 735  
Qy 213 AsnSerTyrSerAlaLeuSerLysGlnLeuGlnLeuAlaSerThrAsnAsnGlyGlnPhe 232  
Db 736 AATAGTTGGGGAGACTTTCCACTGCAATTCAA-----GAGTCTAACCAAGAGGCGCTTT 789  
Qy 233 GluSerProValValLeuLeuAsnAlaGlnAsnGlnArgValThrIleThrAsnValAsp 252  
Db 790 GCTAGTCCAAATCAACTGCAAAAGAGCTAATGGTTCCAAATTCAGTGTG-----TACGAT 843  
Qy 253 AlaGlyValValThrSerAsnLeuAlaLeuLeuAsnArg-----AsnAsn 268  
Db 844 GTGAGTATATTAATCCCTATCATAGCTCTCATGGTGTATAGATGCGACCTCCACCATCG 903  
Qy 269 MetAlaAlaMetAspAspValProMetThrGlnSerPhe 282  
Db 904 TCACAGTTGTTTTCGAGAACTATCCAAATAGTGCAAAATTTT 945

RESULT 5

US-09-785-921A-1  
; Sequence 1, Application US/09785921A  
; Patent No. US2002009434A1  
; GENERAL INFORMATION:  
; APPLICANT: Keener, William K.  
; APPLICANT: Ward, Thomas E.  
; TITLE OF INVENTION: SELECTIVE DESTRUCTION OF CELLS INFECTED WITH HUMAN  
; TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS  
; FILE REFERENCE: LIT-PI-529  
; CURRENT APPLICATION NUMBER: US/09/785,921A  
; CURRENT FILING DATE: 2001-02-15  
; NUMBER OF SEQ ID NOS: 17  
; SEQ ID NO 1  
; LENGTH: 1698  
; TYPE: DNA  
; ORGANISM: Ricinus communis  
US-09-785-921A-1

```
Alignment Scores:
Pred. No.: 4,8e-45 Length: 1698
Score: 423.00 Matches: 94
Percent Similarity: 60.24% Conservative: 59
Best Local Similarity: 37.01% Mismatches: 85
Query Match: 29.54% Indels: 16
DB: 9 Gaps: 7

US-09-905-247-1 (1-289) x US-09-785-921A-1 (1-1698)
QY 25 ValSerPheArgLeuSerGlyAlaThrSerSerSerTyrGlyValPheIleSerAsnLeu 44
Db 97 ATAAACTTTTACCACAGCGGGTCCACTGTGCAAGCTACACAACTTTATCAGAGCTGT 156
QY 45 ArgLysAlaLeuProAsnGluArgLysLeu---TyrAspIleProLeuLeu-----Arg 61
Db 157 CGCGGTGCTTTAACTGCTGATGTCACCAATGTCATATGTCGCTACCATATACCAAGCTGTGCAACAGCA 216
QY 62 SerSerLeuProGlySerGlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAspGlu 81
Db 217 GTTGTGTTTGCCTATAACCAACCGTTTATTTAGTTGAACCTCTCAATCATGCAAGCTT 276
QY 82 ThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGlyAsp 101
Db 277 TCTGTATACATAGCCCTGATGTCCCAATGTCATATGTCGCTACCGTCTGGAAT 336
QY 102 ThrSerTyrPhePhe-----AsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPhe 119
Db 337 AGCGCATATTTCTTCATCCTGACATCAGGAAGATGCAAGCAATCACTCATCTTTTC 396
QY 120 LysAspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThr 139
Db 397 ACTGATGTTCAAAATCGATATACATTCGCTTGTGTGTAATATATGATAGACTTGAACA 456
QY 140 AlaAlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIle 159
Db 457 CTGCTGTAATCTGAGAGAAATATCGAGTTGGGAAATGGTCCACTAGAGAGGCTATC 516
QY 160 ThrThrLeuPheTyrTyrAsnAlaAsn-----SerAlaAlaSerAlaLeu 174
Db 517 TCAGCGCTTTATTTATACAGTACTGGTGGCACTCAGCTTCCAACTCTGCCTCTCTTT 576
QY 175 MetValLeuIleGlnSerThrSerGluAlaAlaArgTyrLysPheIleGluGlnIle 194
Db 577 ATAATTTGCATCCAAATGATTTTCAGAGCAGCAAGATTCCAATATATTGAGGAGAAATG 636
QY 195 GlyLysArgVal-----AspLysThrPheLeuProSerLeuAlaIleIleSerLeuGlu 212
Db 637 CGCAGGAAATTAGGTACACCGGAGATCTGCACAGATCTCTAGCGTAATATACACTTGAG 696
QY 213 AsnSerTrpSerAlaLeuSerLysGlnIleAlaSerThrAsnGlyGlnPhe 232
Db 697 AATAGTTGGGGAGACTTTTCCACTGCAATTCAA-----GAGTCTAACCAAGGAGCCTTT 750
QY 233 GluSerProValValLeuIleAsnAlaGlnAsnGlnArgValThrIleThrAsnValAsp 252
Db 751 GCTAGTCCAAATTCACCTGCAAGAGCGTAATGTTTCCAAATTCAGTGTG-----TACGAT 804
QY 253 AlaGlyValValThrSerAsnIleAlaLeuLeuLeuAsnArg 266
Db 805 GTGAGTATATTAACTCCCTATCATAGCTCTCATGGTGTATAGA 846

RESULT 6
US-10-618-560-1
; Sequence 1, Application US/10618560
; Publication No. US20040048784A1
; GENERAL INFORMATION:
; APPLICANT: Keener, William K.
; APPLICANT: Ward, Thomas E.
; TITLE OF INVENTION: SELECTIVE DESTRUCTION OF CELLS INFECTED WITH HUMAN
; FILE REFERENCE: IMMUNODEFICIENCY VIRUS
```





; PRIOR APPLICATION NUMBER: US 60/045,148		
; PRIOR FILING DATE: 1997-10-30		
; NUMBER OF SEQ ID NOS: 142		
; SOFTWARE: PatentIn Ver. 2.0		
; SEQ ID NO 7		
; LENGTH: 1855		
; TYPE: DNA		
; ORGANISM: Artificial Sequence		
; FEATURE:		
; OTHER INFORMATION: pap-218 insert		
US-10-394-511-7		
Alignment Scores:		
Pred. No.:	5,58e-45	Length: 1855
Score:	423.00	Matches: 94
Percent Similarity:	60.24%	Conservative: 59
Best Local Similarity:	37.01%	Mismatches: 85
Query Match:	29.54%	Indels: 16
DB:	16	Gaps: 7
US-09-905-247-1 (1-289) x US-10-394-511-7 (1-1855)		
QY	25 ValSerPheArgLeuSerGlyAlaThrSerSerTyrGlyValPheIleSerAsnLeu	44
Db	136 ATAAACTTTACACAGCGGGTCCACTGTCCAAAGCTACACAACTTTATCAGAGCTGTT	195
QY	45 ArgLysAlaLeuProAsnGluArgLysLeu---TyrAspIleProLeuLeu-----Arg	61
Db	196 CCGCGTCTGTTTAAACAACCTGGAGCTGATGTGAGACATGATATACCAAGTGTGCCAAACAGA	255
QY	62 SerSerLeuProGlySerGlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAspGlu	81
Db	256 GTTGCTTTGGCTATAAACCAACGCGTTTATTATTAGTTGAACCTCTCAAATCATGACAGCTT	315
QY	82 ThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGlyAsp	101
Db	316 TCTGTTCATTAGCGCTGGATGTCCCAATGCATATGTGTGGCTACCGTGTGGAAAT	375
QY	102 ThrSerTyrPhePhe-----AsnGlyAlaSerAlaThrGluAlaLysTyrValPhe	119
Db	376 AGCGCATATTTCTTCATCCTGCACAAATCAGGAAGATCGAAGCAATCACTCATCTTTTC	435
QY	120 LysAspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThr	139
Db	436 ACTGATGTTCAAATTCGATATACATTCGCCCTTTGGTGTGAATATATAGACTTGAACAA	495
QY	140 AlaAlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIle	155
Db	496 CTTGCTGTGAATCTCAGAGAAATATCGAGTTGGAAATGGTCCACTAGAGAGGCTATC	555
QY	160 ThrThrLeuPheTyrTyrAsnAlaAsn-----SerAlaAlaSerAlaLeu	174
Db	556 TCAGCGCTTTATTATTACAGTACTGGTGGCACTTCGCTTCGCTTCGCTTCGCTTCCTTT	619
QY	175 MetValLeuIleGlnSerThrSerGluAlaAlaArgTyrLysPheIleGluGlnGlnIle	194
Db	616 ATAATTGTGATCCAAATGATTTTCAGACGCAAGATTCCAATATATTAGGGAGAAATG	675
QY	195 GlyLysArgVal-----AspLysThrPheLeuProSerLeuAlaIleSerIleuGlu	212
Db	676 CGCACGAGAATTAGGTACAAACCGAGATCTCGACCAGATCTCTAGCGTAATTTACACTT	735
QY	213 AsnSerTyrSerAlaLeuSerLysGlnIleAlaSerThrAsnAsnGlyGlnPhe	233
Db	736 AATAGTGTGGGGAGACTTTCCCACTGCAATTCAA-----GAGTCTAACCAAGGAGCTTT	789
QY	233 GluSerProValValLeuIleAsnAlaGlnAsnGlnArgValThrIleThrAsnValAsp	255
Db	790 GCTAGTCCCAATTCACCTGCAGAAAGCGTAATGGTTCCAAATTCAGTGTG-----TAGCAT	845
QY	253 AlaGlyValValThrSerAsnIleAlaLeuLeuLeuAsnArg	266
Db	844 GTGAGTATATTAATCCCTATCATAGCTCTCATGGTGTATAGA	895

Best Local Similarity:	37.01%	Mismatches:	85
Query Match:	29.54%	Indels:	16
DB:	16	Gaps:	7

US-09-905-247-1 (1-289) x US-10-394-511-5 (1-1855)

QY	25	valserPheArgLeuSerGlyAlaThrSerSerSerTyrGlyValPheIleSerAsnLeu	44
	:	:::::	:::::
Db	136	ATAAATCTTTACCAAGCGGGTGCCACTGTGCAAAGCTACACAACCTTTATCAGAGCTGTT	195
QY	45	ArgLysAlaLeuProAsnGluArgLysLeu---TyrAspileProLeuLeu-----Arg	61
	:	:::::	:::::
Db	196	CGGGTCGTTTAACAACCTGGAGCTGATGTGAGACATGATATACCAGTGTGCCAAAACAGA	255
QY	62	SerSerLeuProGlySerGlnArgTyrAlaLeulleHisLeuThrAsnTyrAlaaspGlu	81
	:	:::::	:::::
Db	256	GTTGGTTTGGCTATPAACCAACCGGTTTTATTAGTTGAACCTCTCAAATCATGACGAGCTT	315
QY	82	ThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGlyAsp	101
	:	:::::	:::::
Db	316	TCTGTTACATTAGCGCTGGATGTCAACAATGCATATGTTGGTCGCTACCGTCTGGAAT	375
QY	102	ThrSerTyrPhePhe-----AsnGlyAlaSerAlaThrGluAlaLaLysTyrValPhe	119
	:	:::::	:::::
Db	376	AGCCGATATTCTTTTCATCTCTGACAATCGGAAGATGCAGAAGCAATCACTCATCTTTTC	435
QY	120	LysAspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThr	139
	:	:::::	:::::
Db	436	ACTGATGTTCAAAATCGATATATACATTCGCCCTTTGGTGTATATGATAGACTTGACAA	495
QY	140	AlaAlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIle	159
	:	:::::	:::::
Db	496	CTTGCTGTAACTCTGAGAGAAAAATATCGAGTTGGGAAATGGTCCACTAGAGAGGCTATC	555
QY	160	ThrThrLeuPheTyrTyrAsnAlaasn-----SerAlaalaSerAlaLeu	174
	:	:::::	:::::
Db	556	TCAGCGCTTTATATFACAGTACTGGTGCACTCACGTTCCAACTCTGCTCGTCCCTTT	615
QY	175	MetValLeulleGlnSerThrSerGluAlaAlaArgTyrLysPheileGlnGlnIle	194
	:	:::::	:::::
Db	616	ATAATTGCATCCAAATGATTTTCAGAGCAGCAAGATTCCAATATATATTAGGGAGAAATG	675
QY	195	GlyLysArgVal-----AspLysThrPheLeuProSerLeuAlaIleIleSerLeuGlu	212
	:	:::::	:::::
Db	676	GCACAGAAATTAGGTACUACCCGGAGATCTGCACCATCTCCAGGATCTTACGTTGAG	735
QY	213	AsnSerTrpSerAlaLeuSerLysGlnIleGlnIleAlaSerThrAsnAsnGlyGlnPhe	232
	:	:::::	:::::
Db	736	AATAGTTGGGAGAGACTTTCCACTGCAATTTCAA-----GAGTCTAACCAAGGAGCCTTT	789
QY	233	GluSerProValValLeulleAsnAlaGlnAsnGlnArgValThrIleThrAsnValAsp	252
	:	:::::	:::::
Db	790	GCTAGTCCAATTCACCTGCAAAAGACGTAATAGTTGCCAAATTCAGTGTG-----TACGAT	843
QY	253	AlaGlyValValThrSerAsnIleAlaLeuLeuLeuAsnArg	266
	:	:::::	:::::
Db	844	GTGAGTATATTAAATCCCTATCATAGCTCTCATGGGTATAGA	885

RESULT 11  
US-10-394-511-7  
; Sequence 7, Application US/10394511  
; Publication No. US20040009551A1  
; GENERAL INFORMATION:  
; APPLICANT: Borgford, Thor  
; TITLE OF INVENTION: Ricin-like Toxin Variants for Treatment of Cancer,  
; TITLE OF INVENTION: Viral or Parasitic Infections  
; FILE REFERENCE: 10447-27  
; CURRENT APPLICATION NUMBER: US/10/394,511  
; CURRENT FILING DATE: 2003-03-24  
; PRIOR APPLICATION NUMBER: US 09/403,752  
; PRIOR FILING DATE: 1999-10-29  
; PRIOR APPLICATION NUMBER: US 60/063,715  
; PRIOR FILING DATE: 1997-10-29

RESULT 12  
US-10-394-511-9  
; Sequence 9, Application US/10394511  
; Publication No. US20040009551A1  
; GENERAL INFORMATION:  
; APPLICANT: Borgford, Thor  
; TITLE OF INVENTION: Ricin-Like Toxin Variants for Treatment of Cancer,  
; FILE REFERENCE: 10447-27  
; CURRENT APPLICATION NUMBER: US/10/394,511  
; CURRENT FILING DATE: 2003-03-24  
; PRIOR APPLICATION NUMBER: US 09/403,752  
; PRIOR FILING DATE: 1999-10-29  
; PRIOR APPLICATION NUMBER: US 60/063,715  
; PRIOR FILING DATE: 1997-10-29  
; PRIOR APPLICATION NUMBER: US 60/045,148  
; PRIOR FILING DATE: 1997-10-30  
; NUMBER OF SEQ ID NOS: 142  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 9  
; LENGTH: 1855  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: pAP-220 insert  
US-10-394-511-9

Alignment Scores:  
Pred. No.:        5,58e-45        Length:        1855  
Score:            423.00        Matches:      94  
Percent Similarity: 60.24%     Conservative:   59  
Best Local Similarity: 37.01%   Mismatches:   85  
Query Match:       29.54%     Indels:        16  
DB:                16          Gaps:          7

US-09-905-247-1 (1-289) x US-10-394-511-9 (1-1855)

Qy	25	ValSerPheArgLeuSerGlyAlaThrSerSerTyrGlyValPheIleSerAsnLeu	44
Db	136	ATAAACTTTACCAAGCGGGTGCCACTGTGCAAAAGTCACAAATTCATCAGAGCTGT	195
Qy	45	ArgLysAlaLeuProAsnGluArgLysLeu---TyrAspIleProLeuLeu-----Arg	61
Db	196	CGCGGTCTGTTAACAACACTGGAGCTGATGTGACATCATATACCAGTGTGCCAACACAGA	255
Qy	62	SerSerLeuProGlySerGlnArgTyrrAlaLeuIleHisLeuThrAsnTyrrAlaAspGlu	81
Db	256	GTTGGTTTGCCTATTAACCAACCGGTTATTATTAGTTGAACCTCTCAATCATCGACAGCTT	315
Qy	82	ThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrrAtqAlagIlyAsp	101
Db	316	TCTGTTTACATTAGCGTGGATGTCCCAATGCATATGTGTGGCTACCGGTCTGGAAAT	375
Qy	102	ThrSerTyrrPhePhe-----AsnGlyAlaSerAlaThrGluAlaLaLysTyrrValPhe	119
Db	376	AGCGCATATTCTTTCATCCTGCACAAATCAGGAAGATGCAAGCAATCACATCTCTTTC	435
Qy	120	LysAspAlaMetArgLysValThrLeuProTyrrSerGlyAsnTyrrGluArgLeuGlnThr	139
Db	436	ACTGATGTTCAAATTCGATATACATTCGCCCTTTGGTGGTAAATTATGATAGACTTGAACA	495
Qy	140	AlaAlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIle	159
Db	496	CTTGCTGGTAATCTCAGAGAAATATCGAGTTGGGAATATGGTCCACTAGAGAGGCTATC	555
Qy	160	ThrThrLeuPheTyrrTyrrAsnAlaAsn-----SerAlaAlaSerAlaLeu	174
Db	556	TCAGCGCTTTATTATTACAGTACTCGTGGCACTCAGCTTCCAACCTCGTGGCTCGTTCCTT	615
Qy	175	MetValLeuIleGlnSerThrSerGluAlaAlaArgTyrrLysPheIleGluGlnGlnIle	194
Db	616	ATAAATTGTCATCAAATGATTTCAGAACGACAGATTCCAAATGATTATTTGGGGAGAAATG	675

Qy 195 GlyLysArgVal-----AspLysThrPheLeuProSerLeuAlaIleIleSerLeuGlu 211  
Db 676 CGCACGAGAATTAGGTACCAACCGGAGATCTGCACAGATCTCTAGCGTAATTACACTTGAG 735  
Qy 213 AsnSerTrpSerAlaLeuSerLysGlnIleGlnIleAlaSerThrAsnAsnGlyGlnPhe 232  
Db 736 AATAGTTGGGGAGACTTTCACCTGCAATTCAA-----GAGCTTAACCAAGGACGCTTT 789  
Qy 233 GluSerProValValLeuIleAsnAlaGlnAsnGlnArgValThrIleThrAsnValAsp 252  
Db 790 GCTAGTCCAAATTCAACTGCAAAAGACGTAATGTTCCAAATTCAGNGTG-----TACGAT 843  
Qy 253 AlaGlyValValThrSerAsnIleAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 266  
Db 844 GTGAGTATATTAATCCCTATCATAGCTCTCATGCTGTATAGA 885

RESULT 13  
US-10-394-511-13  
; Sequence 13, Application US/10394511  
; Publication No. US20040009551A1  
; GENERAL INFORMATION:  
; APPLICANT: Borgford, Thor  
; TITLE OF INVENTION: Ricin-Like Toxin Variants for Treatment of Cancer,  
; FILE REFERENCE: 10447-27  
; CURRENT APPLICATION NUMBER: US/10/394,511  
; CURRENT FILING DATE: 2003-03-24  
; PRIOR APPLICATION NUMBER: US 09/403,752  
; PRIOR FILING DATE: 1999-10-29  
; PRIOR APPLICATION NUMBER: US 60/063,715  
; PRIOR FILING DATE: 1997-10-29  
; PRIOR APPLICATION NUMBER: US 60/045,148  
; PRIOR FILING DATE: 1997-10-30  
; NUMBER OF SEQ ID NOS: 142  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 13  
; LENGTH: 1855  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: pAP-224 insert  
US-10-394-511-13

Alignment Scores:  
Pred. No.:        5,58e-45        Length:        1855  
Score:            423.00        Matches:      94  
Percent Similarity: 60.24%     Conservative:   59  
Best Local Similarity: 37.01%   Mismatches:   85  
Query Match:       29.54%     Indels:        16  
DB:                16          Gaps:          7

US-09-905-247-1 (1-289) x US-10-394-511-13 (1-1855)

Qy	25	ValSerPheArgLeuSerGlyAlaThrSerSerTyrGlyValPheIleSerAsnLeu	44
Db	136	ATAAACTTTACCAAGCGGGTGCCACTGTGCAAAAGTCACAAATTCATCAGAGCTGT	195
Qy	45	ArgLysAlaLeuProAsnGluArgLysLeu---TyrAspIleProLeuLeu-----Arg	61
Db	196	CGCGGTCTGTTAACAACACTGGAGCTGATGTGACATCATATACCAGTGTGCCAACACAGA	255
Qy	62	SerSerLeuProGlySerGlnArgTyrrAlaLeuIleHisLeuThrAsnTyrrAlaAspGlu	81
Db	256	GTTGGTTTGCCTATTAACCAACCGGTTATTATTAGTTGAACCTCTCAATCATCGACAGCTT	315
Qy	82	ThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrrAtqAlagIlyAsp	101
Db	316	TCTGTTTACATTAGCGTGGATGTCCCAATGCATATGTGTGGCTACCGGTCTGGAAAT	375
Qy	102	ThrSerTyrrPhePhe-----AsnGlyAlaSerAlaThrGluAlaLaLysTyrrValPhe	119
Db	376	AGCGCATATTCTTTCATCCTGCACAAATCAGGAAGATGCAAGCAATCACATCTCTTTC	435
Qy	120	LysAspAlaMetArgLysValThrLeuProTyrrSerGlyAsnTyrrGluArgLeuGlnThr	139
Db	436	ACTGATGTTCAAATTCGATATACATTCGCCCTTTGGTGGTAAATTATGATAGACTTGAACA	495
Qy	140	AlaAlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIle	159
Db	496	CTTGCTGGTAATCTCAGAGAAATATCGAGTTGGGAATATGGTCCACTAGAGAGGCTATC	555
Qy	160	ThrThrLeuPheTyrrTyrrAsnAlaAsn-----SerAlaAlaSerAlaLeu	174
Db	556	TCAGCGCTTTATTATTACAGTACTCGTGGCACTCAGCTTCCAACCTCGTGGCTCGTTCCTT	615
Qy	175	MetValLeuIleGlnSerThrSerGluAlaAlaArgTyrrLysPheIleGluGlnGlnIle	194
Db	616	ATAAATTGTCATCAAATGATTTCAGAACGACAGATTCCAAATGATTATTTGGGGAGAAATG	675

Qy 195 GlyLysArgVal-----AspLysThrPheLeuProSerLeuAlaIleIleSerLeuGlu 211  
Db 676 CGCACGAGAATTAGGTACCAACCGGAGATCTGCACAGATCTCTAGCGTAATTACACTTGAG 735  
Qy 213 AsnSerTrpSerAlaLeuSerLysGlnIleGlnIleAlaSerThrAsnAsnGlyGlnPhe 232  
Db 736 AATAGTTGGGGAGACTTTCACCTGCAATTCAA-----GAGCTTAACCAAGGACGCTTT 789  
Qy 233 GluSerProValValLeuIleAsnAlaGlnAsnGlnArgValThrIleThrAsnValAsp 252  
Db 790 GCTAGTCCAAATTCAACTGCAAAAGACGTAATGTTCCAAATTCAGNGTG-----TACGAT 843  
Qy 253 AlaGlyValValThrSerAsnIleAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 266  
Db 844 GTGAGTATATTAATCCCTATCATAGCTCTCATGCTGTATAGA 885



```
Qy 120 LysAspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThr 139
Db 436 ACTGATGTTCAAAATCGATATATACATTCGCCCTTGGTGAATATATAGACTTTGAACAA 495
Qy 140 AlaAlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIle 159
Db 496 CTTGCTGTGTAATCTGAGAGAAATATCGAGTTGGGAATGCTCCACATAGAGGAGCTATC 555
Qy 160 ThrThrLeuPheTyrTyrAsnAlaAsn-----SerAlaAlaSerAlaLeu 174
Db 556 TCAGCGCTTATTATATACAGTACTGGTGGCACTCAGCTTCCAACTCTCGCTCGTTCCTTT 615
Qy 175 MetValLeuIleGlnSerThrSerGluAlaAlaArgTyrLysPheIleGluGlnIle 194
Db 616 ATAATTTTCATCCAAATGATTTTCAGACGACGAAAGATTCCAAATATATTGGGGAGAAATG 675
Qy 195 GlyLysArgVal-----AspLysThrPheLeuProSerLeuAlaIleSerLeuGlu 212
Db 676 CGCAGAGAAATAGTACACCGGAGATCTGCACAGATCTTAGCGTAAATACACTTGAG 735
Qy 213 AsnSerTrpSerAlaLeuSerLysGlnIleGlnIleAlaSerThrAsnAsnGlyGlnPhe 232
Db 736 AATAGTTGGGGAGACTTTCCACTGCAATTCAA-----GAGTCTAACCAAGGAGCCCTTT 789
Qy 233 GluSerProValValLeuIleAsnAlaGlnAsnGlnArgValThrIleThrAsnValAsp 252
Db 790 GCTAGTCCAAATCAACTGAAAGAGCTAATGGTTCCAAATTCAGTGTG-----TACGAT 843
Qy 253 AlaGlyValValThrSerAsnIleAlaLeuLeuLeuAsnArg 266
Db 844 GTGAGTATATTAATCCCTATCATAGTCTCATGGTGTATAGA 885
```

## RESULT 14

```
US-10-394-511-15
; Sequence 15, Application US/10394511
; Publication No. US20040009551A1
; GENERAL INFORMATION:
; APPLICANT: Borgford, Thor
; TITLE OF INVENTION: Ricin-Like Toxin Variants for Treatment of Cancer,
; FILE REFERENCE: 10447-27
; CURRENT APPLICATION NUMBER: US/10/394,511
; PRIOR FILING DATE: 2003-03-24
; PRIOR APPLICATION NUMBER: US 09/403,752
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 60/063,715
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: US 60/045,148
; PRIOR FILING DATE: 1997-10-30
; NUMBER OF SEQ ID NOS: 142
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 15
; LENGTH: 1855
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pap-226 insert
US-10-394-511-15
```

```
Alignment Scores:
Pred. No.: 5,58e-45 Length: 1855
Score: 423.00 Matches: 94
Percent Similarity: 60.24% Conservative: 59
Best Local Similarity: 37.01% Mismatches: 85
Query Match: 29.54% Indels: 16
DB: 16 Gaps: 7
```

US-09-905-247-1 (1-289) x US-10-394-511-15 (1-1855)

```
Qy 25 ValSerPheArgLeuSerGlyAlaThrSerSerTyrGlyValPheIleSerAsnLeu 44
Db 136 ATAACCTTTACCAAGGGGTGCCACTGTGCAAAAGCTACACAACTTTATCAGAGCTGTT 195
```

```
Qy 45 ArgLysAlaLeuProAsnGluArgLysLeu---TyrAspIleProLeuLeu-----Arg 61
Db 196 CGCGGTGTTTAAACAACCTGAGCTGATGTGAGACATGATATACCAAGTGTTCACCAACAGA 255
Qy 62 SerSerLeuProGlySerGlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAspGlu 81
Db 256 GTTGGTTTGGCTATAAACCAACGCTTTATTATTAGTTGAACCTCTCAAAATCATCGAGAGCTT 315
Qy 82 ThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGlyAsp 101
Db 316 TCTGTTACATTAGCGCTGGATGTCACCAATGCATATGTGTCGCTACCGTGGCTGGAAT 375
Qy 102 ThrSerTyrPhePhe-----AsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPhe 119
Db 376 AGCGCATATTCTTTTCATCTCGACAAATCAGGAAGATGCAGAAACAATCACTCATCTTTTC 435
Qy 120 LysAspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThr 139
Db 436 ACTGATGTTCAAAATCGATATACATTCGCCCTTGGTGAATATATGATAGACTTTGAACAA 495
Qy 140 AlaAlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIle 159
Db 496 CTTGCTGTGTAATCTGAGAGAAATATCGAGTTGGGAATGTCACATAGAGGAGCTATC 555
Qy 160 ThrThrLeuPheTyrTyrAsnAlaAsn-----SerAlaAlaSerAlaLeu 174
Db 556 TCAGCGCTTATTATATACAGTACTGGTGGCACTCAGCTTCCAACTCTCGCTCGTTCCTTT 615
Qy 175 MetValLeuIleGlnSerThrSerGluAlaAlaArgTyrLysPheIleGluGlnIle 194
Db 616 ATAATTTTCATCCAAATGATTTTCAGACGACGAAAGATTCCAAATATATTGGGGAGAAATG 675
Qy 195 GlyLysArgVal-----AspLysThrPheLeuProSerLeuAlaIleSerLeuGlu 212
Db 676 CGCAGAGAAATAGTACACCGGAGATCTGCACAGATCTTAGCGTAAATACACTTGAG 735
Qy 213 AsnSerTrpSerAlaLeuSerLysGlnIleGlnIleAlaSerThrAsnAsnGlyGlnPhe 232
Db 736 AATAGTTGGGGAGACTTTCCACTGCAATTCAA-----GAGTCTAACCAAGGAGCCCTTT 789
Qy 233 GluSerProValValLeuIleAsnAlaGlnAsnGlnArgValThrIleThrAsnValAsp 252
Db 790 GCTAGTCCAAATCAACTGAAAGAGCTAATGGTTCCAAATTCAGTGTG-----TACGAT 843
Qy 253 AlaGlyValValThrSerAsnIleAlaLeuLeuLeuAsnArg 266
Db 844 GTGAGTATATTAATCCCTATCATAGTCTCATGGTGTATAGA 885
```

## RESULT 15

```
US-10-394-511-17
; Sequence 17, Application US/10394511
; Publication No. US20040009551A1
; GENERAL INFORMATION:
; APPLICANT: Borgford, Thor
; TITLE OF INVENTION: Ricin-Like Toxin Variants for Treatment of Cancer,
; FILE REFERENCE: 10447-27
; CURRENT APPLICATION NUMBER: US/10/394,511
; PRIOR FILING DATE: 2003-03-24
; PRIOR APPLICATION NUMBER: US 09/403,752
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 60/063,715
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: US 60/045,148
; PRIOR FILING DATE: 1997-10-30
; NUMBER OF SEQ ID NOS: 142
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 17
; LENGTH: 1855
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pap-226 insert
US-10-394-511-17
```



OTHER INFORMATION: PAP-228 insert  
US-10-394-511-17

## Alignment Scores:

Pred. No.:	5,58e-45	Length:	1855
Score:	423.00	Matches:	94
Percent Similarity:	60.24%	Conservative:	59
Best Local Similarity:	37.01%	Mismatches:	85
Query Match:	29.54%	Indels:	16
DB:	16	Gaps:	7

US-09-905-247-1 (1-289) x US-10-394-511-17 (1-1855)

```
QY 25 ValSerPheArgLeuSerGlyAlaThrSerSerTyrGlyValPheIleSerAsnLeu 44
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 136 ATAACCTTTACCACAGCGGTGCCCTGTGCNAAGCTACACAACTTTATCAGAGCTGTT 195
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 45 ArgLysAlaLeuProAsnGluArgLysLeu---TyrAspIleProLeuLeu-----Arg 61
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 196 CGCGGTCGTTTAAACAACCTGGAGCTGATGTGAGACATGATATACCACTGTTCACAAACAGA 255
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 62 SerSerLeuProGlySerGlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAspGlu 81
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 256 GTTGGTTTCCCTATAACCAACCGGTTTATTATTAGTTGAACCTCTCAATCATGCAGAGCTT 315
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 82 ThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGlyAsp 101
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 316 TCTGTTACATTAGCGCTGGATGTCACCAATGCATATGTGTCGCTACCGTCTGGAAAT 375
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 102 ThrSerTyrPhePhe-----AsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPhe 119
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 376 AGCGCATATTTCTTCATCTGCACATCAGGAAGATGCAGAGCAATCACTCATCTTTTC 435
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 120 LysAspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThr 139
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 436 ACTGATGTTCAAAATCGATATACATTCGGCTTTGGTGGTAATTATGATAGACTTGAACAA 495
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 140 AlaAlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIle 159
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 496 CTTCGTGGTAATCTGAGAGAAATATCGAGTTGGGAATGGTCCACTAGAGGAGGTATC 555
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 160 ThrThrLeuPheTyrTyrAsnAlaAsn-----SerAlaAlaSerAlaLeu 174
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 556 TCAGCGCTTTATTATTACAGTACTGTGTGGCACTCAGCTTCCAACCTCTGGCTCGTTCCCTTT 615
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 175 MetValLeuIleGlnSerThrSerGluAlaAlaArgTyrLysPheIleGluGlnGlnIle 194
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 616 ATAAITTCATCCAAATGATTTCCAGAGCAGCAAGATTCCAATATATTAGGAGGAGAAATG 675
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 195 GlyLysArgVal-----AspLysThrPheLeuProSerLeuAlaIleIleSerLeuGlu 212
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 676 CGCAGAGAAATTAGGTACAAACCGGAGATCTGCACACAGATCTTAGCGTAATTACACTTGAG 735
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 213 AsnSerTrpSerAlaLeuSerLysGlnIleGlnIleAlaSerThrAsnAsnGlyGlnPhe 232
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 736 AATAGTTGGGGGAGACTTTCCACTGCAATTCAA-----GAGTCTAAACCAAGGAGCCTTT 789
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 233 GluSerProValIleAlaLeuAsnAlaGlnAsnGlnArgValThrIleThrAsnValAsp 252
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 790 GCTAGTCCAATTCAACTGAAAGAGCGTAATGGTTCCAATTCAGTGTG-----TACGAT 843
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 253 AlaGlyValValThrSerAsnIleAlaLeuLeuLeuAsnArg 266
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 844 GTGAGTATATTAAATCCCTATCATAGCTCTCATGGGTGTATAGA 885
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

Search completed: September 14, 2004, 23:07:40  
Job time : 463 secs

